

REVISTA DE ESTUDOS DA LINGUAGEM

Faculdade de Letras da UFMG

ISSN

Impresso: 0104-0588

On-line: 2237-2083

V.25 - Nº 3



REVISTA DE ESTUDOS DA LINGUAGEM

Universidade Federal de Minas Gerais

REITOR: Jaime Arturo Ramirez

VICE-REITORA: Sandra Regina Goulart Almeida

Faculdade de Letras:

DIRETORA: Graciela Inés Ravetti de Gómez

VICE-DIRETOR: Rui Rothe-Neves

Editora-chefe

Heliana Ribeiro de Mello (UFMG)

Marcus Maia (UFRJ)

Editores-associados

Adelande Pereira Ferraz (UFMG)

Gustavo Ximenes Cunha (UFMG)

Maria Cândida Trindade Costa de Seabra (UFMG)

Revisão e Normalização

Heliana Ribeiro de Mello

Alda Lopes Durães Ribeiro

Ricardo Alkimim

Editoração eletrônica

Alda Lopes Durães Ribeiro

Úrsula Francine Massula

Secretária

Úrsula Francine Massula

Capa

Elson Rezende de Melo

REVISTA DE ESTUDOS DA LINGUAGEM, v.1 - 1992 - Belo Horizonte, MG,
Faculdade de Letras da UFMG

Histórico:

1992 ano 1, n.1 (jul/dez)

1993 ano 2, n.2 (jan/jun)

1994 Publicação interrompida

1995 ano 4, n.3 (jan/jun); ano 4, n.3, v.2 (jul/dez)

1996 ano 5, n.4, v.1 (jan/jun); ano 5, n.4, v.2; ano 5, n. esp.

1997 ano 6, n.5, v.1 (jan/jun)

Nova Numeração:

1997 v.6, n.2 (jul/dez)

1998 v.7, n.1 (jan/jun)

1998 v.7, n.2 (jul/dez)

1. Linguagem - Periódicos I. Faculdade de Letras da UFMG, Ed.

CDD: 401.05

ISSN: Impresso: 0104-0588

On-line: 2237-2083

REVISTA DE ESTUDOS DA LINGUAGEM

V. 25 - Nº 3- jul.-set. 2017

Indexadores

DOAJ (Directory of Open Access Journals) [Sweden]

EBSCO [USA]

JournalSeek [USA]

Linguistics & Language Behavior Abstracts [USA]

MLA Bibliography [USA]

WorldCat / OCLC (Online Computer Library Center) [USA]

MIAR (Matriu d'Informació per a l'Anàlisi de Revistes) [Spain]

Latindex [Mexico]

Sindex (Scientific Indexing Services) [USA]

CSA - Linguistics and Language Behavior Abstract

Portal CAPES

DRJI (Directory of Research Journals Indexing) [India]

OAJI (Open Academic Journals Index) [Russian Federation]

REDIB (Red Iberoamericana de Innovación y Conocimiento Científico) [Spain]



REVISTA DE ESTUDOS DA LINGUAGEM

Editora-chefe

Heliana Ribeiro de Mello (UFMG)

Editores Associados

Aderlande Pereira Ferraz (UFMG)

Gustavo Ximenes Cunha (UFMG)

Maria Cândida Trindade Costa de Seabra (UFMG)

Comissão Científica

Aderlande Pereira Ferraz (UFMG)

Alessandro Panunzi (Univ. degli Studi Firenze, Itália)

Alina Maria S. M. Villalva (Univ. de Lisboa)

Ana Lúcia de Paula Müller (USP)

Augusto Soares da Silva (UCP, Braga)

Beth Brait (PUCSP / USP)

Carmen Lucia Barreto Matzenauer (UCPEL)

César Nardelli Cambraia (UFMG)

Charlotte C. Galves (UNICAMP)

Cristina Name (UFJF)

Deise Prina Dutra (UFMG)

Diana Luz Pessoa de Barros (USP/Mackenzie-SP)

Dylia Lysardo-Dias (UFSJ)

Edwiges Morato (UNICAMP)

Emília Mendes Lopes (UFMG)

Esmeralda V. Negrão (USP)

Gabriel de Ávila Othero (UFRGS)

Gerardo Augusto Lorenzino (Temple University)

Gláucia Muniz Proença Lara (UFMG)

Hanna Batoréo (Universidade Aberta, Lisboa)

Heliana Ribeiro de Mello (UFMG)

Heronides Moura (UFSC)

Hilário Bohn (UCPEL)

Hugo Mari (PUC Minas)

Ida Lúcia Machado (UFMG)

Ieda Maria Alves (USP)

Ivã Carlos Lopes (USP)

Jairo Nunes (USP)

João Antônio de Moraes (UFRJ)

João Miguel M. da Costa (Univ. Nova, Lisboa)

João Queiroz (UFJF)

João Saramago (Univ. de Lisboa)

John Robert Schmitz (UNICAMP)

José Borges Neto (UFPR)

Kanavillil Rajagopalan (UNICAMP)

Leo Wetzels (Free University of Amsterdam)

Leonel Figueiredo de Alencar (UFC)

Lodenir Becker Kamopp (UFRGS)

Lorenzo Teixeira Vitral (UFMG)

Luiz Amaral (Univ. of Massachusetts Amherst)

Luiz Carlos Cagliari (UNESP-Araraquara)

Luiz Carlos Travaglia (UFU)

Marcelo Barra Ferreira (USP)

Márcia Maria Cañado Lima (UFMG)

Márcio Leitão (UFP)

Marcus Maia (UFRJ)

Maria Antonieta Amarante M. Cohen (UFMG)

Maria Bernadete Marques Abaurre (UNICAMP)

Maria Cecília Camargo Magalhães (UFRJ)

Maria Cecília Magalhães Mollica (UFRJ)

Maria Cândida Trindade C. de Seabra (UFMG)

Maria Cristina Figueiredo Silva (UFPR)

Maria do Carmo Viegas (UFMG)

Maria Luíza Braga (PUC-RJ)

Maria Marta P. Scherre (UnB)

Milton do Nascimento (PUC-Minas)

Mônica Santos de Souza Melo (UFV)

Philippe Martin (Université Paris 7)

Rafael Nonato (Museu Nacional / UFRJ)

Raquel Meister Ko. Freitag (UFS)

Roberto de Almeida (Concordia University)

Ronice Müller de Quadros (UFSC)

Ronald Beline (USP)

Rove Chishman (UNISINOS)

Sanderléia Longhin-Thomazi (UNESP)

Sérgio de Moura Menuzzi (UFRGS)

Seung-Hwa Lee (UFMG)

Sírio Possenti (UNICAMP)

Thaís Cristófaros Alves da Silva (UFMG)

Tony Berber Sardinha (PUC-SP)

Ubiratã Kickhöfel Alves (UFRGS)

Vander Viana (University of Stirling, UK)

Vanice Gomes de Medeiros (UFF)

Vera Lúcia Lopes Cristóvão (UEL)

Vera Lúcia Menezes de O. e Paiva (UFMG)

Vilson J. Leffa (UCPEL)

Sumário / Contents

Panorama da Linguística Experimental no Brasil <i>An Overview of Experimental Linguistics in Brazil</i> Marcus Maia	951
A escolha de software e hardware na psicolinguística: revisão e opinião <i>The choice of software and hardware in psycholinguistics: review and opinion</i> Thiago Oliveira da Motta Sampaio	971
Maxakalí Nasality and Field Recording with Earbud Microphony <i>Nasalidade no Maxakalí e gravação em campo com microfonia de fone de ouvido</i> Andrew Ira Nevins Mário André Coelho da Silva	1011
Social affective variations in Brazilian Portuguese: a perceptual and acoustic analysis <i>Variações socioexpressivas no português do Brasil: análise perceptiva e acústica</i> Albert Rilliard João Antônio de Moraes	1043
Perception of emotional prosody: investigating the relation between the discrete and dimensional approaches to emotions <i>Percepção da prosódia emocional: investigando a relação entre as abordagens discreta e dimensional das emoções</i> Wellington da Silva Plínio Almeida Barbosa	1075

Prosódia de enunciados declarativos e interrogativos totais nas variedades de Salvador, Fortaleza e Rio de Janeiro <i>Prosody of declarative and interrogative modalities in Salvador, Fortaleza and Rio de Janeiro varieties</i> Leticia Rebollo Couto Carolina Gomes da Silva Luma da Silva Miranda	1105
Desambiguação de sentenças na interface fonologia-sintaxe: resultados de um estudo de compreensão <i>Sentence disambiguation in the phonology-syntax interface: comprehensive study results</i> Melanie Campilongo Angelo Raquel Santana Santos	1143
Representation and Processing of the Inflected Infinitive in Brazilian Portuguese: an eye-tracking study <i>Representação e processamento dos infinitivos flexionados em português brasileiro: um estudo com rastreamento ocular</i> Marcello Modesto Marcus Maia	1183
On the asymmetry between subject and object relative clauses in discourse context <i>Assimetria no custo de processamento de relativas de sujeito e de objeto em contextos discursivos</i> Renê Forster Leticia Maria Sicuro Corrêa	1225
A aceitabilidade da anáfora logofórica em português brasileiro <i>The acceptability of logophoric anaphor in Brazilian Portuguese</i> Flávia Gonçalves Calaça de Souza Rosana Costa de Oliveira Judithe Genuíno Henrique	1255

Efeitos de distância linear e marcação no processamento da concordância verbal variável no PB <i>Linear distance and markedness effects in variable subject-verb agreement processing in BP</i> Mercedes Marcilese Erica dos Santos Rodrigues Marina Rosa Ana Augusto Késsia da Silva Henrique	1291
Agreement effects of gender and number in pronominal coreference processing in Brazilian Portuguese <i>Efeitos de concordância de gênero e número no processamento da coreferência pronominal em português brasileiro</i> Michele Calil dos Santos Alves	1327
A behavioral study to investigate the processing routes of grammatical gender in Brazilian Portuguese <i>Um estudo comportamental para investigar as rotas do processamento do gênero gramatical do português brasileiro</i> Natália C. A. de Resende Mailce Borges Mota	1367
A influência da referencialidade no processamento de orações relativas em português brasileiro <i>The influence of referentiality on relative clause processing in Brazilian Portuguese</i> Gitanna Brito Bezerra Márcio Martins Leitão Lorrane da Silva Neves Medeiros	1397
An ERP study of kind-denoting nouns in the subject position in Brazilian Portuguese <i>Um estudo neurofisiológico de nomes que denotam tipo na posição de sujeito em português brasileiro</i> Aniela Improta França Maurício Cagy Antonio Fernando Catelli Infantosi	1433

Referência genérica em SNs singulares: uma abordagem cognitivista experimental <i>Generic reference in singular Nps: a cognitive grammar experimental approach</i> Diogo Pinheiro Lilian Ferrari Maria Clara Pimenta Joabe de Souza Márcia Viegas Wanderson Lucas Ohenes	1463
Um estudo experimental do processamento de metáforas do português brasileiro <i>An experimental study on Brazilian Portuguese metaphor processing</i> Antonio João Carvalho Ribeiro Adiel Queiroz Ricci	1501
A técnica de ERP no processamento de sentenças de crianças: uma revisão <i>The ERP technique in children's sentence processing: a review</i> Marília Uchôa Cavalcanti Lott de Moraes Costa	1537
Aquisição de estruturas possessivas inalienáveis: o caso dos nomes de parte do corpo em inglês americano e português brasileiro <i>Acquisition of inalienable possessive structures: the case of body-part names in American English and Brazilian Portuguese</i> Fernanda Mendes	1567
On the acquisition of distributivity in Yudja <i>Sobre a aquisição da distributividade em Yudja</i> Suzi Lima	1613
O processamento do numeral reduplicado sohoji-sohoji em Karajá: uma averiguação de ERP durante a compreensão de sentenças <i>The processing of the reduplicated numeral sohoji-sohoji in Karajá: An ERP investigation during sentence comprehension</i> Cristiane Oliveira da Silva	1647

Are bilingualism effects on the L1 byproducts of implicit knowledge? Evidence from two experimental tasks <i>Efeitos do bilinguismo sobre a L1 são produtos de conhecimento implícito? Evidências de duas tarefas experimentais</i> Ricardo Augusto de Souza Cândido Samuel Fonseca de Oliveira	1685
Morphological priming resists language and modality switching in late Dutch-Brazilian Portuguese bilinguals <i>Priming morfológico resiste em face à alternância entre línguas e modalidades com bilingues tardios de holandês e português brasileiro</i> Marije Soto Aline Gesualdi Manhães	1717
O teste de Cloze como instrumento de medida da proficiência em leitura: fatores linguísticos e não linguísticos <i>Cloze test as an instrument to measure reading proficiency: linguistic and non-linguistic factors</i> Kátia Nazareth Moura de Abreu Daniela Cid de Garcia Katharine de Freitas P. N. A. da Hora Cristiane Ramos de Souza	1767
O movimento ocular na leitura realizada por revisores de textos <i>Revisers' eye movement in reading</i> Délia Ribeiro Leite José Olímpio de Magalhães	1801

Panorama da Linguística Experimental no Brasil

An Overview of Experimental Linguistics in Brazil

Marcus Maia

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
CNPq

A linguagem é uma faculdade da mente humana e, como tal, é implementada por processos cognitivos complexos. Os estudos linguísticos baseados na inspeção de *corpora* podem analisar os produtos destes processos, o que tem, sem dúvida, permitido descobertas importantes sobre as línguas. Entretanto, naturalmente, esses estudos estão restritos aos dados que inspecionam, não sendo capazes, por definição, de examinar os processos que geraram tais produtos. Se restrita à análise de *corpora*, portanto, a pesquisa linguística ficaria impossibilitada de testar hipóteses sobre questões linguísticas fundamentais, como por exemplo a gramaticalidade ou a aceitabilidade das expressões linguísticas.

Como discutimos em Maia (2015), para estudar tais processos, a gramática gerativa propôs, desde a sua criação, em meados dos anos 50, que julgamentos sobre a boa formação de frases fossem coletados. Buscando uma caracterização da capacidade gerativa da linguagem, Noam Chomsky propôs que a gramaticalidade das frases poderia ser testada através de julgamentos intuitivos obtidos informalmente pelo linguista consigo próprio e/ou com alguns falantes. Embora o próprio Chomsky tenha avaliado, posteriormente, que a confiabilidade exclusiva da teoria em julgamentos intuitivos informais deveria ser substituída por critérios mais rigorosos,¹ o fato é que a coleta de julgamentos informais

¹ “Como qualquer outra pessoa, também não gosto de me fiar em intuições. Devemos substituir as intuições por critérios rigorosos, assim que pudermos”. (cf. CHOMSKY discussion in HILL, 1962, p. 177)

introspectivos sobre a aceitabilidade ou sobre a gramaticalidade de frases tem sido a principal ferramenta de avaliação de dados em linguística gerativa há várias décadas. Sua proposição inicial foi verdadeiramente revolucionária em um momento da história da ciência dominada pelo behaviorismo antimentalista. Por meio desses testes informais, a linguística gerativa pôde operar uma mudança importante no foco dos estudos lingüísticos, que pôde ir além das análises exaustivas dos *corpora* para oferecer uma modelagem inicial dos processos cognitivos subjacentes ao conhecimento da linguagem.

Desde esses dias pioneiros, os experimentos lingüísticos têm percorrido um longo caminho no sentido de adotarem-se critérios mais rigorosos, utilizando metodologias e tarefas engenhosas, análises estatísticas apropriadas e até mesmo os equipamentos sofisticados dos laboratórios de psicolinguística, podendo oferecer evidências independentes, além das intuições iniciais do linguista, para explorar os processos cognitivos da linguagem.

O presente número da Revista de Estudos da Linguagem – RELIN reúne uma amostra significativa de estudos experimentais realizados no Brasil. São 24 artigos, representando as principais técnicas experimentais em uso no mundo para explorar o conhecimento da linguagem, flagrando processos lingüísticos *off-line* e *on-line*, tais como os testes de ranqueamento em escala, o julgamento controlado de gramaticalidade, o julgamento de valor de verdade, o teste *Cloze*, o paradigma de *priming*, a decisão lexical, a leitura automonitorada, a tarefa *maze*, a extração de potenciais evocados, o rastreamento ocular, entre outros. Esses estudos investigam diferentes aspectos de questões fonéticas, fonológicas, morfológicas, sintáticas, semânticas e pragmáticas, bem como questões de interface entre vários desses subsistemas do conhecimento lingüístico, oferecendo um panorama amplo da Linguística Experimental no país.

Antes de passarmos à apresentação de cada um dos artigos, queremos destacar ainda um fato que é também digno de nota a respeito do presente volume da RELIN: três dos trabalhos aqui reunidos tomam como objeto de suas pesquisas experimentais dados obtidos junto a falantes de línguas indígenas brasileiras, a saber, a língua Maxakalí (família Jê), a língua Yudjá (Tronco Tupi) e a língua Karajá (Tronco Macro-Jê). Como se sabe, foi apenas muito recentemente que a pesquisa de campo com línguas indígenas começou a beneficiar-se do método experimental. Já era tempo que a revolução cognitivista chegasse à linguística de campo e esse

número da RELIN situa-se entre os primeiros a publicar estudos dessa natureza no Brasil. Este novo empreendimento, que está sendo chamado, justamente, de pesquisa de campo experimental, está enfrentando o desafio de reunir dimensões cruciais da linguística, como as diferentes teorias, os métodos psicolinguísticos e os procedimentos etnográficos de campo, a fim de tentar explorar processos gramaticais que dificilmente poderiam ser descobertos apenas com base em análise de *corpora*.

Feito este destaque, passemos à apresentação dos artigos. Após a resenha do primeiro artigo, de **natureza metodológica**, seguem-se cinco artigos que estudam **questões fonéticas e fonológicas ou de interface entre a fonologia e a sintaxe**. Apresentam-se, na sequência, artigos que discutem **questões de processamento e de representação gramatical** e um grupo de artigos que tem como objeto **a relação de concordância**. Os próximos quatro artigos debruçam-se sobre o tema da **referência nominal**. Seguem-se três artigos que estudam experimentalmente aspectos da **aquisição da linguagem**, o último deles investigando questões relativas à **distributividade** em uma língua indígena, tema também abordado pelo artigo seguinte, com falantes adultos de outra língua indígena. O **processamento bilíngue** é tema de dois artigos. A revista encerra com dois artigos de viés aplicado sobre o **tema da leitura**.

Em um momento em que a possibilidade de experimentação vem atraindo pesquisadores de praticamente todas as subáreas da Linguística, o artigo de Sampaio que abre o volume, intitulado **A escolha de software e hardware na psicolinguística: revisão e opinião** oferece orientações seguras da psicolinguística, principalmente para os recém-chegados ao campo fascinante dos métodos experimentais. Em um texto bastante didático, Sampaio discorre sobre os diferentes programas usados na implementação de experimentos, sugerindo uma série de critérios a serem considerados na sua escolha, além de oferecer informações práticas que podem poupar tempo para os linguistas decididos a explorar metodologias experimentais.

Cinco estudos neste número investigam fenômenos fonéticos, fonológicos ou na interface fonologia-sintaxe. Em **Maxakalí nasality and field recording with earbud microphony**, Nevins e Silva iniciam revendo dados da língua indígena Maxakalí (família Jê) que demonstram que esta língua tem processos fonológicos tais como pré-nasalização, nasalização em coda, harmonia nasal dentro da sílaba, entre núcleo e rima, entre rima e *onset* e de sílaba a sílaba, que oferecem

desafios interessantes para as teorias fonológicas. Após discutirem as técnicas mais tradicionalmente utilizadas no estudo da nasalidade, os autores apresentam a sua proposta inovadora de fonologia de campo experimental: a técnica de gravação com microfonia de fone de ouvido, que é, então, demonstrada em relação a fenômenos de nasalidade vocálica e consonantal, empréstimos, entre outros fenômenos na língua Maxakalí. Ao final, os autores fazem ainda uma interessante prospecção de áreas que a técnica de microfonia de fone de ouvido poderia vir a impactar positivamente, incluindo questões pedagógicas.

No artigo **Social affective variations in Brazilian Portuguese: a perceptual and acoustic analysis**, Rilliard e Moraes partem de um experimento de escolha livre com 22 ouvintes brasileiros para analisar parâmetros prosódicos de duração, intensidade e entonação, tomando como quadros de referência para a análise, teorias de expressividade na fala e de uso simbólico da voz em interações verbais. Os autores descrevem agrupamentos de expressões, rotulando-as como “pergunta neutra”, “admiração”, “arrogância”, “autoridade”, “desprezo”, “obviedade”, “incerteza”, “ironia”, “irritação”, “pisando em ovos”, “polidez”, “sedução”, “sinceridade”, “surpresa”, “descrédito”, entre outras expressões. Os resultados do estudo permitem aos autores reagrupar dezessete expressões em três grupos principais de valência positiva, assertividade e dubitatividade. Os autores comparam ainda os dados brasileiros com dados obtidos com franceses e japoneses, para concluir que os dados brasileiros apresentam maior acurácia na singularização de expressões únicas, em comparação com as duas outras línguas.

Também pesquisando a percepção de emoções na fala, Silva e Barbosa discutem no artigo **Perception of emotional prosody: investigating the relation between the discrete and dimensional approaches to emotions**, as relações entre a percepção de emoções básicas e suas dimensões emocionais na fala, conduzindo dois experimentos de percepção de enunciados em uma língua estrangeira sobre a qual os participantes não tinham qualquer conhecimento. Com base nesses experimentos, os autores puderam correlacionar as dimensões emocionais ativação, justiça, valência, motivação e envolvimento com as emoções básicas alegria, raiva, tristeza e calma. Em seguida, as emoções básicas e dimensões emocionais foram correlacionadas com parâmetros acústicos, confirmando que os ouvintes, de fato, utilizam ao menos parcialmente traços acústicos para julgar a expressividade das emoções

e dimensões emocionais. Ao final, os autores sugerem que esse tipo de investigação tem potencial para contribuir com pesquisas sobre a síntese da fala e sobre o reconhecimento automático de emoções.

Em **Prosódia de enunciados declarativos e interrogativos totais nas variedades de Salvador, Fortaleza e Rio de Janeiro**, Rebollo Couto, Silva e Miranda gravam enunciados declarativos e interrogativos, elicitando-os a partir de imagens. As autoras avaliam que esses dados de fala experimental, assim coletados, seriam mais próximos dos dados de leitura do que dos dados provenientes de fala atuada, em que o falante já conhece o texto, tendo praticado previamente a sua elocução. Os participantes da pesquisa são adultos de escolaridade média, de ambos os sexos, originários de Salvador, Fortaleza e Rio de Janeiro, havendo os dados obtidos no teste, sido analisados acusticamente no programa PRAAT em termos de parâmetros acústicos de F0, duração e intensidade das vogais pré-tônicas, tônicas e pós-tônicas. Os resultados das análises dos dados obtidos no experimento são comparados a resultados de pesquisas similares nas três capitais, permitindo aos autores demonstrar a existência de diferenças prosódicas entre as modalidades declarativa e interrogativa testadas, bem como variações dialetais.

Angelo e Santos apresentam em **Desambiguação de sentenças na interface fonologia-sintaxe** os resultados de um experimento do tipo *oral sentence/picture matching* (pareamento de sentença auditiva e imagem), com o objetivo de analisar como pistas de duração na desambiguação de sentenças do tipo SN1-Verbo SN2-Atributo influenciam a interpretação dos ouvintes. O artigo situa-se em uma linha de pesquisa muito produtiva que vem testando o princípio de processamento *Late Closure* (Aposição Local) em diferentes construções em várias línguas. As autoras apresentaram frases como “A mãe encontrou a filha suada” em contextos que favoreciam ora a interpretação do tipo *early closure* (a mãe estava suada) ou a interpretação *late closure* (a filha estava suada), manipulando também no *input* auditivo a duração do trecho entre a sílaba final do objeto até a sílaba inicial do atributo, enquanto os participantes visualizavam imagens que deveriam escolher, de acordo com a interpretação das frases que ouviam, levando em conta ainda gradações no alongamento. Os resultados obtidos demonstram a operação do princípio *Late Closure*, (cf. FRAZIER, 1979), em português brasileiro, corroborando o trabalho de Magalhães e Maia (2006), segundo o qual uma preferência pelo *default* local ocorre na ausência de pistas prosódicas.

Diversos estudos neste volume apresentam resultados experimentais sobre a compreensão adulta de construções sintáticas que, além de fornecerem evidências relevantes para teorias de processamento, também estabelecem diálogo com questões de representação gramatical. Em **Representation and Processing of the Inflected Infinitive in Brazilian Portuguese: an eye-tracking study**, Modesto e Maia apresentam estudo de rastreamento ocular que indica menor custo significativamente para o processamento da correferência entre antecedentes acessíveis na frase e a categoria vazia PRO do que nos casos em que a interpretação do PRO é arbitrária. Em linha com a literatura psicolinguística relativa aos chamados modelos de “dois estágios”, que têm estabelecido como *default* no processamento sintático a computação intrassentencial “syntax first”, precedendo a extrassentencial, os achados de Maia e Modesto são também pertinentes para contribuir com dados empíricos para debate corrente na teoria sintática e na descrição da gramática do português brasileiro, constituindo, neste sentido, empreendimento realizado não só no quadro de referência da especialidade do processamento de frases, mas também no quadro da especialidade relativamente mais recente da sintaxe experimental. A literatura de natureza teórica, descritiva e também sociolinguística sobre o infinitivo flexionado em português é revisada criteriosamente na primeira parte do artigo, sugerindo-se que “os contextos sintáticos de uso da flexão não finita no Brasil atual e a interpretação de tais estruturas são questões de grande interesse científico, considerando-se a situação sociopolítica do país”. O ponto fulcral do debate sobre a representação do infinitivo flexionado em português situa, de um lado, a análise de Rodrigues e Hornstein (2013), que afirma que as construções com infinitivo flexionado em português são configurações de controle não obrigatório e, de outro, a análise de Modesto (2010, 2016a), que argumenta serem os sujeitos de orações não finitas flexionadas normalmente interpretados como sujeitos controlados em PB. Os resultados do experimento, que indicam que as construções com PRO controlado em frases com verbos infinitivos flexionados não apenas são psicologicamente reais, mas são, de fato, a opção preferida em uma comparação estrita com o PRO de interpretação arbitrária, servem de fundamentação empírica para a conclusão em favor da proposta de Modesto (2010, 2016a).

No artigo **On the asymmetry between subject and object relative clauses in discourse context**, Forster e Corrêa também

examinam, através de experimento de rastreamento ocular, questão de processamento com relevância para a teoria gramatical minimalista. Realizado no âmbito do quadro maior do projeto de pesquisa da segunda autora, intitulado “Processamento e aquisição da linguagem sob ótica minimalista: extensão e comparação de modelos”, o estudo toma como objeto de investigação a assimetria existente no processamento de orações relativas de sujeito *vis à vis* o processamento de orações relativas de objeto, manipulando como uma das variáveis independentes do experimento o contexto narrativo em que a oração relativa é apresentada, que poderia favorecer o seu processamento como relativa de sujeito ou de objeto. O experimento tinha, portanto, um *design* 2x2, contendo 40 narrativas (20 favorecendo a relativa de objeto, outras 20 a relativa de sujeito), distribuídas em quadrado latino e apresentadas a 41 participantes, para leitura, ranqueamento e interpretação. Os resultados obtidos indicam maior custo para as relativas de objeto do que para as relativas de sujeito, independentemente do contexto discursivo. Destaque-se que os autores avaliam seus resultados não só em termos de teorias de processamento, trazendo à discussão o debate clássico em psicolinguística entre modelos de dois estágios *versus* estágio único, mas também, interessadamente, em termos de desenvolvimentos recentes da linguística gerativa que, como já tem apontado o grupo de pesquisa dos autores (LAPAL), vem minimizando as diferenças que, tradicionalmente, tem colocado em oposição os modelos de um e dois estágios.

O artigo **A aceitabilidade da anáfora logofórica em português brasileiro**, de Souza, Oliveira e Henrique, oferece um exemplo paradigmático de empreendimento em sintaxe experimental. Considerando explicitamente em seu texto que a intuição do linguista poderia não ser suficiente para permitir avaliar adequadamente a aceitabilidade das construções logofóricas em PB, tema do artigo, as autoras desenvolvem um estudo controlado, ainda que não cronométrico, de julgamento de aceitabilidade para testar sentenças com anáfora (e.g. O genro de Mauro cortou José e **ele mesmo** com a faca na cozinha) e com pronome (e.g. O genro de Mauro cortou José e **ele** com a faca na cozinha) logofóricos. Segundo avaliam no texto, as intuições das autoras, ponto de partida do estudo, indicavam que ambas as construções pudessem ser aceitáveis em PB, o que se revelou inadequado, como demonstrado pelos resultados obtidos no teste, em que a aceitabilidade dos pronomes logofóricos não atingiu níveis de significância. As autoras discutem

seus resultados com base em teorias de representação como a Teoria da Ligação (CHOMSKY, 1981) e, principalmente, com base na Teoria da Reflexividade de Reinhart e Reuland (1993). Além da interessante discussão empreendida pelas autoras, fica comprovada no estudo a importância da máxima frequentemente mencionada na especialidade da sintaxe experimental: as intuições devem ser ponto de partida, mas não ponto de chegada das pesquisas.

Um tema em foco em estudos recentes em psicolinguística e em sintaxe experimental é o processamento da concordância que, segundo Phillips, Wagers e Lau (2011), tem constituído um desafio para os psicolinguistas interessados em pesquisar a compreensão *on-line* das construções gramaticais. Phillips e colegas avaliam que, nas construções que envolvem concordância, o processador sintático parece mais suscetível à falibilidade, ao contrário do que ocorre em outras construções (e.g. efeitos de ilha, dependências entre antecedentes e categorias vazias, princípios de *Binding*), em que o processador demonstra “sensibilidade *on-line* impressionante” (cf. p. 149). O tema da concordância requer, portanto, muita pesquisa para que possamos vir a compreender melhor como os falantes “navegam representações linguísticas complexas em tempo real”,² como avaliam os autores, após um levantamento abrangente da literatura. No presente número da RELIN, três artigos se propõem a contribuir, através de experimentos, para avançar nosso entendimento da relação de concordância.

Em **Efeitos de distância linear e marcação no processamento da concordância verbal variável em PB**, Marcilese, Rodrigues, Augusto e Henrique estudam a computação das relações de concordância entre sujeito e verbo, manipulando a distância linear entre esses elementos, em um teste do tipo *maze task* ou labirinto, em que os participantes devem escolher, entre duas opções, como continuar a leitura de uma frase. Interessantemente, os autores partem de achados da sociolinguística variacionista, que têm indicado, sistematicamente, a relevância da distância linear entre sujeito e verbo na realização em *corpora* da concordância verbal em português brasileiro. Após rever criteriosamente esses estudos e também a literatura pertinente em psicolinguística, as

² “...the parser’s highly selective vulnerability to interference and “grammatical illusions” provides a valuable tool for understanding how speakers encode and navigate complex linguistic representations in real time” (cf. PHILLIPS; WAGERS; LAU, 2011, p. 147.)

autoras se propõem a estudar, para além dos produtos encontrados nos *corpora*, o processo cognitivo de concordância, manipulando a distância linear como uma das variáveis independentes do experimento, ao lado dos fatores número do verbo e número do sujeito. Seus resultados indicam não só a existência de efeito de distância linear no processamento da concordância, mas também “apontam para a relevância da marcação na atuação dos mecanismos de verificação de traços”. O artigo é concluído reconhecendo-se a importância de empreendimentos interdisciplinares, como o que foi realizado no estudo em tela, buscando um “diálogo explícito” entre variação e processamento, para se conseguir avançar no entendimento de fenômenos linguísticos complexos como a concordância.

O tema da marcação da concordância em PB é também investigado em **Agreement effects of gender and number in pronominal coreference processing in Brazilian Portuguese**, em que Alves reporta um estudo de rastreamento ocular que tem como objetivo específico o de explorar como e quando as condições estruturais impostas pelo Princípio B da Teoria da Ligação (CHOMSKY, 1981) interagem com pistas de concordância, na recuperação mnemônica de antecedentes por pronomes. A autora mede a duração de fixações oculares em áreas críticas de frases apresentadas a 24 participantes, para testar a hipótese de que a redundância na marcação morfológica da concordância em PB levaria os leitores a utilizar tais pistas na recuperação de antecedentes por pronomes. São encontrados efeitos tanto de antecedentes estruturalmente aceitáveis, quanto de antecedentes não aceitáveis estruturalmente, levando a autora a alinhar-se com os achados de Badecker e Straub (2002), contrariando a hipótese de que as condições estruturais atuem como um filtro inicial no processamento. Finalmente, tomando como base seus resultados, a autora também argumenta em favor do modelo *CAM – content addressable memory*, que propõe que informações previamente analisadas possam ser recuperadas a partir de uma busca, em paralelo, de pistas gramaticais, como as oferecidas pela marcação da concordância de número e de gênero em PB.

O terceiro artigo que investiga experimentalmente a relação gramatical de concordância em PB é **A behavioral study to investigate the processing routes of grammatical gender in Brazilian Portuguese**, de Resende e Mota. Diferentemente dos dois estudos anteriores, este artigo focaliza as relações de concordância no âmbito do Sintagma Determinante (DP), para pesquisar possíveis diferenças nas rotas de

processamento da concordância de gênero em duas condições, a saber, artigo e nome (condição 1) e nome e adjetivo (condição 2) em uma tarefa em que se mediram tempos de reação dos participantes nas duas condições, contrastando-se também substantivos masculinos e femininos, transparentes quanto ao seu gênero gramatical (terminações -a e -o) e opacos (outras terminações). Os resultados obtidos indicam a existência de diferenças no processamento da concordância de gênero entre as condições 1 e 2, atribuídas à maior quantidade de informação a ser processada na relação entre o nome e adjetivo do que na relação entre o artigo e o nome. As autoras também encontram efeitos de frequência e de memória, influenciando tanto o processamento de formas transparentes, quanto de formas opacas, o que as leva a concluir em favor dos modelos unitários, não duais, de processamento da linguagem.

O importante tema da referência de nomes é explorado através de angulações teóricas e de metodologias diversas, em quatro artigos, neste número da RELIN, devotado à Linguística Experimental. Em **A influência da referencialidade no processamento de orações relativas em português brasileiro**, Bezerra, Leitão e Medeiros, testam, em estudos de leitura automonitorada e de rastreamento ocular de orações relativas do PB, o Princípio da Referencialidade, apresentado no âmbito do modelo *Construal* (FRAZIER; CLIFTON, 1996), que prediz que orações adjetivas restritivas sejam preferencialmente apostas a núcleos referenciais. Os autores se colocam o desafio de tentar flagrar através de medidas *on-line* efeito de referencialidade operando a tempo de influenciar as decisões do processador sintático. Como, a partir do modelo *Construal*, relações sintáticas secundárias (como as orações relativas), seriam associadas e não apostas (casos restritos à aposição de argumentos, em relações sintáticas primárias), no marcador frasal em construção, o processador seria capaz de avaliar informação semântica e pragmática para informar suas decisões, ao contrário das propostas iniciais do tipo “*syntax-first*” da Teoria do *Garden-Path*. Os autores corroboram em PB as propostas do modelo *Construal* e discutem similaridades e diferenças entre os resultados obtidos nos experimentos de leitura automonitorada e de rastreamento ocular, este último viabilizado graças à cooperação entre os laboratórios LAPROL (UFPB) e LAPEX (UFRJ).

Em **An ERP study of kind-denoting nouns in subject position in Brazilian Portuguese**, França, Maurício e Infantosi colocam em contraste uma abordagem do tipo “*syntax-first*”, inspirada no Programa

Minimalista (cf. CHOMSKY, 1995) e uma abordagem do tipo “semantic-first”, que toma como referência propostas de natureza cognitivista de Fillmore (1982, 2007). Para estabelecer o contraditório entre as duas abordagens os autores tomam como objeto do estudo o tema central da referencialidade, orquestrando de forma bastante lógica e clara, em estudo de extração de potenciais cerebrais evocados, a comparação entre Sintagmas Nominais de denotação específica e Sintagmas Nominais que denotam tipo em um contexto especial de conhecimento de mundo. A hipótese do estudo é a de que, em linha com a abordagem “*syntax-first*”, os nomes que denotam tipo, em posição de sujeito, não deverão influenciar o processamento em medidas iniciais em que o conteúdo semântico do nome não deveria ser acessado. De fato, monitorando 34 participantes através de equipamento de EEG com 20 canais, enquanto estes realizavam julgamentos de gramaticalidade de frases em que as condições experimentais eram contrastadas, tomando-se como controle condições incongruentes que deveriam produzir ERP (N400), os autores puderam concluir que “o pleno acesso ao conteúdo pragmático que está no sujeito só ocorre após o merge do complemento ao verbo”.

O artigo **Referência genérica em SNs singulares: uma abordagem cognitivista experimental** de Pinheiro, Ferrari, Pimenta, Souza, Viegas e Ohenes, também estabelece como objeto de estudo aspectos da referencialidade de nomes, mas adota perspectiva teórica diametralmente oposta à do estudo apresentado no parágrafo acima, tomando como quadro de referência a Gramática Cognitiva (cf. LANGACKER, 1987, 2013, entre outros), que propõe caracterização de base semântica de sintagmas nominais. Os autores postulam, inicialmente, a existência de uma diferença conceptual entre o SN genérico definido (e.g. o gato é voluntarioso) e o SN genérico indefinido com determinante zero (e.g. gato é voluntarioso). Alinhando-se com o que chamam de “guinada empírica” em Linguística Cognitiva” (cf. STEFANOWITSCH, 2011), os autores se propõem, em seguida, a tentar estabelecer experimentalmente o efeito que uma sequência com contraste poderia exercer sobre o SN singular genérico com artigo definido, manipulando a presença e a ausência dessas sequências contrastivas nos estímulos (e.g. A criança gosta de doces e guloseimas vs A criança gosta de doces e guloseimas. Já o adulto tende a ter um paladar mais apurado). Em uma tarefa *off-line* de julgamento de aceitabilidade controlada, utilizando uma escala *Likert* de cinco pontos, os autores testaram 30 participantes para

comparar frases como exemplificadas acima, distribuídas em quadrado latino. Após analisar seus resultados, os autores concluem que podem “confirmar a previsão de que a presença de um segundo SN genérico, com função contrastiva, na mesma sequência textual eleva significativamente a aceitabilidade de enunciados com SN genérico definido”. Tratando-se de experimento *off-line*, este resultado não pode, portanto, ser usado para comparar questões de curso temporal, com possível impacto na arquitetura da representação e do processamento linguísticos, a partir do exame da referencialidade de nomes, como as entretidas nos dois estudos apresentados anteriormente, que utilizam variáveis dependentes *on-line*. Ainda assim, a perspectiva de início de um programa de pesquisa experimental em Linguística Cognitiva, de que esse trabalho é, sem dúvida, um dos precursores no Brasil, tem o potencial de trazer para o exame empírico criterioso da Linguística Experimental questões que tradicionalmente têm fundamentado divergências importantes e, muitas vezes, motivado debates acalorados, em Linguística.

O artigo assinado por Ribeiro e Ricci, intitulado **Um estudo experimental do processamento de metáforas do português brasileiro** também examina uma questão relacionada à referência de nomes, a saber, a computação do sentido literal e não literal ou metafórico. Depois de relatar dois estudos normativos bastante criteriosos, que lhes permitiram estabelecer, com confiança, controles relativos à familiaridade, adequação e convencionalidade das metáforas nominais, os autores reportam um estudo de leitura automonitorada em que foram apresentadas a 66 participantes, em distribuição em quadrado latino, frases alvos precedidas de três tipos de contextualizações, motivando interpretação metafórica, interpretação literal, inclusão em classe. Não tendo observado diferenças significativas nos tempos médios de leitura dos segmentos críticos, nas três condições, os autores argumentam em favor do processamento direto de metáforas familiares, conforme preconizado por Glucksberg e Keysar (1990), habilmente rebatendo o chamado *argumentum ad ignorantiam* (a ausência de evidência não é evidência da ausência), com base no poder do teste estatístico utilizado na análise dos dados.

Três estudos no presente número temático da RELIN investigam experimentalmente a aquisição da linguagem. Em **A técnica de ERP no processamento de sentenças de crianças: uma revisão**, Costa presta uma contribuição para a área de aquisição da linguagem, revisando estudos que têm utilizado a técnica de extração de potenciais cerebrais

evocados por eventos linguísticos, incluindo um estudo seu próprio (Costa, 2015), desenvolvido para sua tese de doutorado. O artigo inicia com uma apresentação bastante didática da técnica EEG-ERP, destacando sua excelente resolução temporal e seu baixo custo, comparativamente a técnicas tais como o MEG, o fMRI, o PET, entre outras. O artigo, então, passa a rever componentes cerebrais frequentemente encontrados em experimentos com adultos, com especial atenção para o N400 e o P600. Passa-se a discutir, em seguida, as razões pelas quais estudos encefalográficos ainda são incipientes em estudos com crianças, revendo-se, na sequência, estudos de segmentação da fala e de processamento de sentenças. Costa conclui seu artigo apresentando o relato de seu próprio experimento, em que testou crianças francesas quanto à estrutura argumental de verbos. Ela relata haver captado uma ativação negativa entre 200-450ms após as crianças terem escutado o verbo que estava sendo manipulado, o que lhe permite estabelecer que as crianças são sensíveis a manipulações na estrutura argumental.

Em **Aquisição de estruturas possessivas inalienáveis: o caso dos nomes de parte do corpo em inglês americano e português brasileiro**, Mendes faz sintaxe experimental ao estudar construções de posse inalienável em PB e em inglês, comparando o desenvolvimento dessas estruturas na criança com as propriedades que as caracterizam na gramática adulta. Após rever a literatura teórica relevante (e.g. Teoria da Ligação, CHOMSKY, 1981, entre vários outros trabalhos teóricos), bem como pesquisas sobre a posse inalienável e sua aquisição, a autora estabelece suas hipóteses e previsões, que serão testadas em uma série de experimentos de julgamento de valor de verdade. Através desses experimentos, testa-se a compreensão de falantes de inglês americano e de PB a respeito da ocorrência de determinantes definidos e pronomes possessivos em construções de posse em que, tanto a leitura alienável, quanto a inalienável, estariam disponíveis. Os resultados obtidos permitem à autora concluir que, embora partindo da mesma gramática universal, a distinção entre as gramáticas do PB e do inglês americano apareceu por volta dos seis anos de idade, quando as crianças falantes de inglês americano passaram a limitar o uso do determinante definido, de acordo com a gramática alvo.

No artigo **On the Acquisition of distributivity in Yudja**, Lima faz semântica formal experimental, estudando a aquisição da distributividade na língua indígena brasileira Yudja, comparativamente a adultos, através

de testes de preferencialidade. Após rever cuidadosamente a literatura relevante tanto em semântica formal, quanto em psicolinguística, Lima estabelece os objetivos da pesquisa: (i) avaliar a disponibilidade da interpretação distributiva encoberta em Yudjá e (ii) explorar a interpretação de frases que marcam a distributividade abertamente na língua, através do processo de reduplicação verbal. Lima descreve, em seguida, as propriedades dos sintagmas nominais e verbais na língua, com especial atenção para a formação da reduplicação verbal, apresentando sua análise teórica sobre a reduplicação e a pluralização de nomes em Yudjá, para então testar experimentalmente a hipótese já formulada em Lima (2008) de que a reduplicação verbal, além de indicar pluralidade, funciona também como um marcador da distributividade aberta na língua. Os resultados dos testes de preferencialidade levados a efeito com onze adultos e dezessete crianças, que a autora considera serem pilotos de futuras pesquisas, indicaram que, enquanto a interpretação dos adultos é influenciada pelo tipo de sintagma nominal, as crianças teriam preferência pela leitura distributiva, independentemente das propriedades do verbo e dos sintagmas nominais.

O trabalho de Silva, intitulado **O processamento do numeral reduplicado *sohoji-sohoji* em Karajá: uma averiguação de ERP durante a compreensão de sentenças** também explora propriedades de distributividade em uma língua indígena brasileira, utilizando muito provavelmente pela primeira vez no Brasil, a técnica de EEG-ERP com uma população indígena. O estudo se desenvolveu no âmbito do projeto guarda-chuva de seu orientador no Programa de Pós-Graduação em Linguística (POSLING-UFRJ) “Pesquisa de campo experimental com línguas indígenas brasileiras”, com o apoio do CNPq. Em janeiro de 2015, uma equipe constituída por professores e doutorandos do POSLING-UFRJ, incluindo os orientadores da tese de Cristiane Oliveira da Silva, os professores Marcus Maia e Aniela Improta França, levaram a campo equipamento de EEG de 64 canais e realizaram testes *in loco* com falantes nativos da língua Karajá, moradores das aldeias Btõiry ou Hãwalò (Ilha do Bananal, TO). Os resultados do teste de *oral sentence-picture matching* (pareamento de sentenças orais e figuras) realizado por 22 participantes falantes nativos de Karajá, enquanto seus sinais cerebrais eram monitorados pelo equipamento de EEG, permitiram à autora identificar componentes N400 e P600, que indicam que o numeral distributivo Karajá *sohoji-sohoji* atuaria como um quantificador

ambíguo, permitindo que leituras evento-distributiva e participante-distributiva fossem licenciadas ou ainda que licenciariam diferentes cenários de verificação para a mesma leitura. A autora conclui, avaliando que a disponibilidade de técnicas cada vez mais sensíveis para aferir o processamento da linguagem pode, de fato, contribuir significativamente para as pesquisas de campo tradicionais e análises formais da linguagem.

Dois estudos neste número da RELIN investigam experimentalmente questões de um tema que vem se tornando cada vez mais presente em Psicolinguística, a saber, o estudo do processamento bilíngue, cujo desafio é o de entender o problema conhecido como *one mind, two languages*. Em **Are bilingualism effects on the L1 byproducts of implicit knowledge? Evidence from two experimental tasks**, Souza e Oliveira estudam efeitos do inglês L2 sobre o desempenho linguístico da L1 português brasileiro, na compreensão de construções resultativas, em dois experimentos psicolinguísticos que elaboram sobre estudo anterior dos autores, utilizando, agora, a tarefa *maze* (*on-line*), além do julgamento imediato de aceitabilidade (*off-line*). A tarefa *maze*, também utilizada no estudo de Marcilese, Rodrigues, Augusto e Henrique, no presente número da RELIN, embora não meça os tempos de leitura dos segmentos em que se dividem as frases, tem sobre a tarefa de leitura automonitorada, a vantagem de não acarretar efeitos *spillover* e nem demandar questões interpretativas, uma vez que força que o leitor escolha ativamente entre duas opções de continuação da frase. Como se sabe construções resultativas, comuns em inglês (e.g. *Samuel wiped the table clean*) não se instanciam no repertório de construções do português. A questão que os autores se colocaram no estudo foi a de verificar a influência da L2 sobre a L1, havendo a tarefa *on-line* se mostrando a única apta a capturar esse efeito, que é analisado pelos autores como um processo implícito.

Soto e Manhães investigam em **Morphological priming resists language and modality switching in late Dutch-Brazilian Portuguese bilinguals**, o processamento morfológico em bilíngues tardios cuja L1 é o holandês e a L2, o português, em uma tarefa de decisão lexical, coletando tanto tempos de reação, quanto dados cerebrais através de EEG. No experimento, monitoraram-se, através de equipamento eletroencefalográfico, 18 bilíngues proficientes (dados de 3 deles seriam posteriormente retirados do estudo), que deveriam decidir se os itens alvos eram palavras, alternando-se itens em português, holandês e, ainda, como controle, itens em húngaro, língua que nenhum dos

participantes conhecia. O estudo pretendia testar a *Superficial Processing Hypothesis* (cf. CLAHSSEN; FELSER, 2006; CLAHSSEN *et al.*, 2010) que prediz processamento morfológico menos eficiente na L2, e o *Revised Hierarchical Model* (cf. KROLL *et al.*, 2010) que, entre outras predições, propõe que haja ativação em paralelo de informações lexicais e sublexicais entre línguas. Os resultados obtidos permitem às autoras estabelecer efeitos morfológicos na L2, nas medidas de EEG, mas não nos tempos de reação, sendo a diferença discutida em termos da maior sensibilidade do N400 a propriedades das raízes do que a palavras inteiras. As autoras concluem que os resultados da alternância entre línguas demonstram *priming* morfológico nas condições em que a direção se dava de L2 para L1, enquanto que, para a direção de L1 para L2, o efeito oposto ocorreu, confirmando-se as previsões feitas pelo Modelo Revisado Hierárquico.

Os dois artigos que encerram o número tratam de questões relacionadas à leitura de textos que, além de examinadas e avaliadas *per se*, também apresentam dimensões aplicadas interessantes. Em **O teste de Cloze como instrumento de medida da proficiência em leitura: fatores linguísticos e não linguísticos**, Abreu, Garcia, da Hora e Souza apresentam inicialmente uma ampla revisão da literatura sobre o teste de *cloze*, uma técnica relativamente simples, que vem sendo usada há pelo menos quatro décadas, em que se solicita aos participantes que completem palavras omitidas em textos. Conforme avaliam as autoras, o teste de Cloze tem aplicações educacionais imediatas, servindo de diagnóstico para avaliar-se a compreensão leitora e permitindo intervenções empiricamente fundamentadas. Após a resenha crítica bem orquestrada, as autoras apresentam o seu próprio experimento baseado na técnica, com o objetivo tanto de avaliar a sensibilidade do teste a fatores linguísticos, quanto não linguísticos. Os testes são organizados em um *design 2x2*, em que se cruzam as variáveis tipo de palavra (funcional ou lexical) e tamanho do texto (curto ou longo), sendo aplicados a 88 alunos, sendo metade do sexto ano do ensino fundamental e metade do primeiro ano do ensino médio. As autoras analisam, então, os resultados obtidos quantitativa e qualitativamente, argumentando, ao final, em favor da aplicação mais ampla do teste de *cloze*, que avaliam ser uma ferramenta versátil, capaz de acessar a proficiência leitora em níveis micro e macroestruturais, tendo o potencial de revelar lacunas no letramento em leitura.

O artigo **O movimento ocular na leitura realizada por revisores de textos**, de Leite e Magalhaes, teve por objetivo estudar, através da técnica de rastreamento ocular, o processamento da leitura em profissionais de revisão de texto, comparativamente a não revisores, entretendo a hipótese de que a leitura dos revisores seria “menos automática, mais controlada e mais detalhada”. Foram submetidos para leitura pelos dois grupos de sujeitos, em tarefa de detecção de erros, 20 parágrafos experimentais, metade com supressão de preposições e metade com anáforas pronominais incorretas, além de 30 textos distratores, parte sem incorreções e parte com diferentes tipos de erros. O índice de detecção de erros nas frases experimentais não indicou diferença significativa entre os dois grupos de sujeitos, ambos de nível superior. Por outro lado, os resultados das medidas de rastreamento ocular permitiram aos autores confirmarem a sua hipótese inicial de que os revisores profissionais teriam uma leitura mais demorada, menos automática, mais detalhada e controlada. Os autores concluem que a pesquisa pode ser informativa tanto para o estudo da leitura em si, quanto para aferir critérios de desempenho de revisores profissionais.

Em conclusão, avaliamos que os 24 artigos reunidos nesse número temático da RELIN apresentam um panorama único, em que se pode apreciar de forma integrada, pela primeira vez no país, uma amostra significativa de pesquisas, que exercitam diferentes angulações teóricas, através das mais praticadas técnicas experimentais no mundo, no exame de um conjunto amplo de questões relevantes para as principais subáreas da Linguística, verdadeiramente atestando a exuberância da Linguística Experimental no Brasil. Esperamos que esses trabalhos sejam fonte de inspiração para novas pesquisas teóricas e experimentais.

Referências

BADECKER, W.; STRAUB, K. The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, Washington, v. 28, p.748-769, 2002. DOI: <http://dx.doi.org/10.1037//0278-7393.28.4.748>.

CHOMSKY, N. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CHOMSKY, N. *The Minimalist Program*. Cambridge, MA: MIT Press, 1995.

CLAHSEN, H.; FELSER, C. Grammatical processing in language learners. *Applied Psycholinguistics*, Cambridge University Press, v. 27, p. 3-42, 2006.

CLAHSEN, H.; FELSER, C.; NEUBAUER, K.; SATO, M; SILVA, R. Morphological structure in Native and Non-Native Language Processing *Language Learning*, Wiley Online Library, v. 60, p. 21-43, 2010.

COSTA, M. U. C. L. M. *Argument Structure in Language Acquisition: an ERP Study*. 2015. 174f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015.

FILLMORE, C. Frame semantics. In: The Linguistic Society of Korea (Ed.). *Linguistics in the morning calm: selected papers from SICOL-1981*. Seoul, Korea: Hashin Publishing Co., 1982. p. 111-137. [Reprinted in: GEERAERTS, D. (Ed.). *Cognitive Linguistics: Basic Reading*. Berlin: Mouton de Gruyter, 2006.]

FILLMORE, C. Valency issues in FrameNet. In: HERBST, T.; GÖTZVOTTELER, K. (Ed.). *Valency: theoretical, descriptive and cognitive issues*. Berlin: Mouton de Gruyter, 2007. p. 129-160.

FRAZIER, L. On comprehending sentences: syntactic parsing strategies. 1979. Tese (Doutorado) – University of Connecticut, 1979. [reproduzida por Indiana University Linguistics Club]

FRAZIER, L.; CLIFTON, C. *Construal*. MIT Press Cambridge, 1996.

GLUCKSBERG, S.; KEYSAR, B. Understanding metaphorical comparisons: beyond similarity. *Psychological Review*, American Psychological Association, v. 97, n. 1, p. 3-18, 1990. DOI: [https://doi.org/10.1016/S1364-6613\(02\)00040-2](https://doi.org/10.1016/S1364-6613(02)00040-2).

HILL, A. (Ed.) *Third Texas conference on problems of linguistic analysis in English*. Second discussion section. Austin: University of Texas, 1962.

KROLL, J. F.; HELL, J. G. van; TOKOWICZ, N.; GREEN, D. The revised hierarchical model: a critical review and assessment. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 13, n. 3, p. 373-381, July, 2010. DOI: <https://doi.org/10.1017/S136672891000009X>.

LANGACKER, R. W. *Foundations of cognitive grammar: theoretical prerequisites*. Stanford: Stanford University Press, 1987.

LANGACKER, R. W. *Essentials of cognitive grammar*. Oxford: Oxford University Press, 2013.

LIMA, Suzi. *A estrutura argumental dos verbos na língua Juruna (Yudja): da formação dos verbos para a análise das estruturas sintáticas*. 2008. Dissertação (Mestrado) – Universidade de São Paulo, São Paulo, 2008.

MAGALHÃES, J. O.; MAIA, M. Pistas prosódicas implícitas na resolução de ambiguidades sintáticas: um caso de adjunção de atributos. *Revista da ABRALIN*, Florianópolis, v.5, n.1-2, p.143-167, 2006.

MAIA, M. A. R. Sintaxe experimental. In: OTHERO, Gabriel; KENEDY, Eduardo. (Org.). *Sintaxe, sintaxes*. São Paulo: Editora Contexto, 2015. p. 51-72.

MODESTO, M. What Brazilian Portuguese says about control: remarks on Boeckx & Hornstein. *Syntax*, Wiley Online Library, v. 13, n. 1, p. 78-96, 2010.

MODESTO, M. Inflected infinitives in Brazilian Portuguese and the theory of control. In: GONÇALVES, A. (Org.). *Complement clauses in Portuguese: syntax and acquisition*. Amsterdam: John Benjamins. 2016a. In press.

PHILLIPS, C.; WAGERS, M. W.; LAU, E. F. Grammatical illusions and selective fallibility. In: RUNNER, J. (Ed.). *Experiments at the interface*. Syntax & Semantics. Bingley, UK: Emerald Publications, 2011. v. 37, p. 153-186.

REINHART, T.; REULAND, E. Reflexivity. *Linguistic Inquiry*, MIT Press, v. 24, n. 4, p. 657-720, 1993.

RODRIGUES, C.; HORNSTEIN, N. Epicene agreement and inflected infinitives when the data is “Under Control”: a reply to Modesto (2010). *Syntax*, Wiley Online Library, v. 16, p. 292-309, 2013.

STEFANOWITSCH, A. Cognitive linguistics as a cognitive science. In: CALLIES, M.; KELLER, W. R.; LOHÖFER, A. (Ed.) *Bidirectionality in the cognitive sciences*. Amsterdam; Philadelphia: John Benjamins, 2011. p. 296-309. DOI: <https://doi.org/10.1075/hcp.32.16ste>.

A escolha de software e hardware na psicolinguística: revisão e opinião

The choice of software and hardware in psycholinguistics: review and opinion

Thiago Oliveira da Motta Sampaio

Language Acquisition, Processing and Syntax Lab - LAPROS

Universidade Estadual de Campinas, Campinas, São Paulo / Brasil

thiagomotta@iel.unicamp.br

Resumo: Nos últimos anos, diversos *softwares* foram criados para auxiliar a elaboração de experimentos em ciências cognitivas. A oferta de *softwares* de simples utilização deveria facilitar o trabalho dos iniciantes, porém, acabou trazendo novos problemas e dúvidas. Que *software* usar? Qual deles é o mais adequado ao meu estudo e por quê? Através de uma revisão sobre computação, linguagem de programação e técnicas de apresentação de estímulos visuais,¹ este artigo pretende fomentar a discussão a respeito (i) dos diversos tipos de *softwares* para estimulação, (ii) da importância de conhecer os detalhes técnicos do *hardware* utilizado e (iii) da compatibilização *hardware-software-método* como uma variável a ser controlada durante o desenvolvimento do protocolo experimental.

Palavras-chave: psicolinguística; ciências cognitivas; linguagens de programação; métodos.

Abstract: In the last few years, several softwares have been designed to help the development of experiments in cognitive sciences. The offer of user friendly software would help beginners in their first tests. However, it brought new problems and questions. Which software should one use? Which one is more adequate for my research and why? The present paper brings a quick and panoramic review on computer science, programming languages and on the presentation of visual stimuli. Through these three topics, I intend to promote a discussion (i) on the main types of software

for stimulation in cognitive sciences, (ii) on the importance of being attentive to the hardware specifications and (iii) on some compatibility issues between software-hardware-method as independent variables in our experiments.

Keywords: psycholinguistics; cognitive sciences; programming languages; methods.

Recebido em: 5 de dezembro de 2016.

Aprovado em: 28 de março de 2017.

1. Introdução

O trabalho a ser apresentado a seguir não irá apresentar uma hipótese para testá-la através do método experimental. Este é um artigo de métodos que tem por objetivo problematizar algumas questões rotineiras de um cientista cognitivo que trabalha com experimentação: o desenvolvimento de experimentos no computador.

Primeiramente é preciso dizer que para elaborar um experimento no computador não é necessário ter um conhecimento muito avançado em computação. Existem diversos softwares especializados nesta tarefa tanto para usuários avançados quanto para os jovens cientistas em formação. Verba para software também não é necessariamente um problema. Para a maior parte dos experimentos, não precisamos desembolsar centenas de dólares para a aquisição de software proprietário, visto que existem diversas opções livres e com uma curva de aprendizagem bastante rápida. Ainda assim, por alguma razão, estes softwares não são tão difundidos entre os pesquisadores da área.

A segunda questão que trago neste artigo é que, apesar da grande oferta de aplicações especializadas, não é aconselhável deixar toda a tarefa de comunicação entre usuário e máquina por conta do software. Ao contrário dos sistemas operacionais mais atuais, os softwares de experimentação não escondem do usuário as configurações impossíveis de serem realizadas pelo seu hardware, nem dão informações precisas sobre o comportamento do hardware durante o teste. Muitas vezes acreditamos ter controle total sobre nossas variáveis sem nos darmos conta de que o computador não está executando exatamente a mesma tarefa que pedimos para ele executar.

Apresentadas minhas questões principais, pretendo iniciar nas próximas seções uma discussão sobre estes dois aspectos. A seção 2, a seguir, problematiza o conhecimento em computação do senso comum. Em seguida, a seção 3 visa esclarecer o que é e como funcionam as linguagens de programação. A seção 4 apresenta diversos tipos de software criados especialmente para experimentação em ciências cognitivas, entre linguagens de programação, toolboxes e softwares com interface gráfica pagos e gratuitos. A seção 5 discursa sobre problemas de controle de experimentos visuais originados do desconhecimento sobre o hardware usado em sua aplicação. Por fim, fecho este artigo com algumas considerações finais.

2. Uma rápida discussão sobre software e hardware

As Ciências da Computação possuem uma regra que parece prever o ritmo do avanço tecnológico: a Lei de Moore (1965). Esta lei demonstra que computadores aumentam exponencialmente sua complexidade, dobrando a capacidade de processamento a cada 2 anos. Se utilizarmos esta regra para olhar para o passado, verificamos que o início da computação teria ocorrido na década de 60, exatamente quando o 1º chip foi inventado.

A partir da lei de Moore, é possível prever que até mesmo usuários entusiastas não conseguem acompanhar o avanço da tecnologia em sua totalidade. Enquanto a quantidade de informação aumenta com o passar dos anos, nossas vidas adquirem mais responsabilidades que nos tomam o tempo necessário para nos atualizar em todas as frentes da tecnologia.

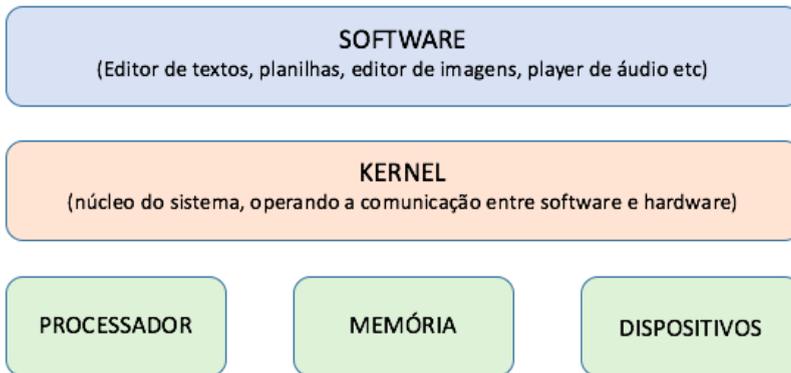
Neste contexto, somos invadidos pela ideia de que as crianças de hoje são nativas digitais e aprendem naturalmente a utilizar tecnologias avançadas que mesmo os entusiastas já não conseguem mais dominar. Porém, em diversas ocasiões, conversas pessoais e até mesmo em uma mesa durante o 3º Encontro de Divulgação Científica e Cultural na Unicamp, foi discutido e constatado que esta é uma meia verdade.

É inegável que existem (i) jovens usuários que realmente aproveitam a maior disponibilidade dos conteúdos digitais para dominarem tecnologias avançadas e se tornarem excelentes programadores. Por outro lado, o que observamos na maioria das vezes é que (ii) os softwares (especialmente os pagos) ficaram cada vez mais acessíveis tanto financeiramente, se tornando mais baratos, quanto na facilidade de utilização (*user friendly*). Isso resulta em uma ilusão de inclusão

tecnológica, na qual muitos usuários menos avançados conseguem, com relativa facilidade, realizar tarefas que seriam bastante complexas para o entusiasta de anos atrás.

Isso pode ser observado com mais clareza ao verificarmos a evolução dos sistemas operacionais dos dispositivos móveis. Cada vez mais populares, estes sistemas que, hoje, quase toda criança tem no bolso, chegaram ao ponto de dividir ou mesmo de substituir algumas funções que, antes, eram realizadas exclusivamente por computadores pessoais caros e inacessíveis. O resultado é que muitas daquelas pessoas que consideramos *experts* em tecnologia da informação possuem nada mais do que um vasto conhecimento sobre como utilizar os diversos softwares disponíveis no mercado. Estes usuários conseguem utilizar diversos tipos de programas computacionais de forma hábil através de sua interface de usuário, no *front-end* (Figura 1). Porém, eles muitas vezes possuem um conhecimento extremamente limitado sobre a comunicação do software com o hardware e da resolução de questões básicas de hardware, o *back-end*.

FIGURA 1 – Kernel (núcleo) do sistema: a ponte entre software e hardware



Nota: O centro de um sistema operacional é o seu núcleo (*kernel*), responsável por servir de ponte entre o software e o hardware de uma máquina. Usuários de interface (*front-end*) geralmente se limitam ao conhecimento do software, não tendo a necessidade de compreender o funcionamento da máquina nos demais níveis. Podemos fazer uma alusão à parte visível de um iceberg, quando a maior parte da rocha está submersa e fora de nosso campo de visão.

Não é difícil encontrar um entusiasta dos anos 90 ou do início do século que tenha lidado com as inúmeras questões de incompatibilidade entre uma nova peça de hardware (ex. placas de som e de vídeo) e a placa mãe de seu computador, ou entre um hardware compatível com sua placa mãe, mas sem a disponibilidade de um controlador para realizar a comunicação entre o sistema operacional e sua nova peça de hardware (*driver*¹).

Quando instalamos o Windows ou o Linux pela primeira vez, eles geram drivers genéricos para que o hardware funcione em suas configurações mínimas. Os drivers específicos de cada peça são procurados após a instalação, para que possamos aproveitar o potencial máximo da máquina. Antigamente, tanto a busca quanto instalação dos drivers eram realizadas de forma manual e, ao passar por estas experiências e buscar por soluções, os usuários acabaram obtendo noções básicas sobre a comunicação entre hardware e software.

Hoje, tanto o hardware, quanto o software, assim como a comunicação entre eles, se tornaram mais eficientes permitindo que o sistema operacional simplesmente esconda muitas das opções de configuração que o hardware não suporta (muito disso devido à presença dos drivers corretos para o hardware instalado), o que evita uma grande parte dos problemas mais básicos que enfrentávamos na década de 90. Além disso, as últimas versões dos sistemas já possuem uma biblioteca dos drivers mais utilizados. Quando os sistemas não possuem o driver de sua peça, eles oferecem a opção de busca automática. Além disso, você sempre pode recuperar o driver na página da fabricante ou nos CDs que acompanham a peça.

Os aparelhos móveis também podem sofrer com a subutilização de suas funções por conta de drivers genéricos. Consideremos a título de exemplo a qualidade das fotos de um smartphone. Ao contrário do iOS, que roda em hardware padronizado e obriga os desenvolvedores de software a usarem os aplicativos e *drivers* oficiais da empresa, o Android precisa se adequar a diversos tipos de hardware de diferentes fabricantes. Por esta razão, cada fabricante possui um tipo de aplicativo de fotos otimizado para o seu aparelho, com diversos filtros que melhoram o desempenho da câmera no pós-processamento. Os aplicativos de terceiros, porém, para não terem o trabalho de criar um aplicativo diferente (ou *drivers*

¹ Software contendo instruções de comunicação entre o sistema operacional e o hardware.

diferentes que aumentariam o tamanho do aplicativo) para cada modelo existente de telefone, geram um driver genérico para acessar diretamente a câmera, o que diminui a qualidade de imagem mesmo nos aplicativos mais usados e famosos.

Conhecer a forma como o computador recebe os inputs pode não ser necessário para todos os usuários desde que tudo funcione aparentemente bem. Por outro lado, este conhecimento pode ser importante para fazermos melhores escolhas de softwares para determinados fins, para resolvermos determinados tipos de problemas técnicos e, especialmente, para quando desenvolvemos nossos experimentos e queremos controlar rigorosamente a forma como o hardware apresenta os estímulos aos participantes e coleta as respostas comportamentais.

O primeiro passo para discutir a comunicação entre software e hardware será o levantamento sobre o que são as linguagens de programação e como elas funcionam. Para isso, me limitarei àquelas que são comumente utilizadas por pesquisadores em ciências cognitivas, o que inclui a psicolinguística, nos principais laboratórios americanos e europeus.

3. O que são as Linguagens de Programação?

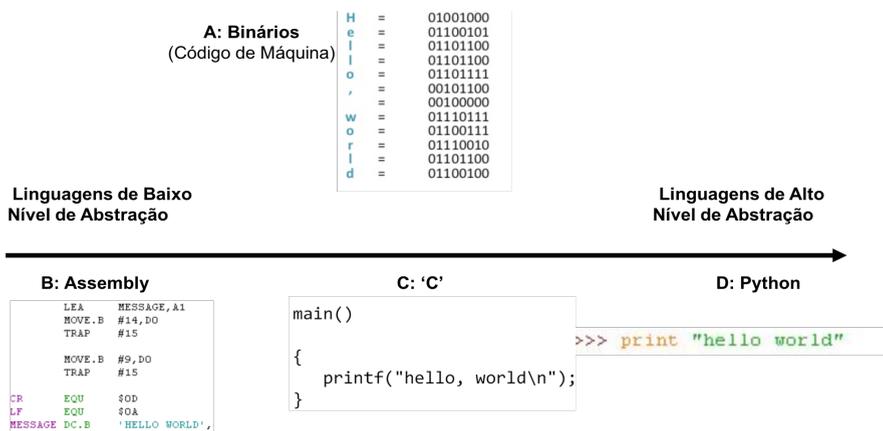
As máquinas clássicas² recebem input em forma de fluxos de corrente elétrica, correspondendo aos números 0 (desativado) e 1 (ativado), o que chamamos de *bits*. Isso nos dá a possibilidade de distinguir até duas informações por vez. Para aumentar o número de informações a serem processadas, os computadores foram desenvolvidos de modo a trabalhar com um conjunto de oito bits (1 *byte*). Agora, cada sequência de 8 bits pode ser relacionada a uma informação diferente. Por exemplo, a representação do número 8 corresponde à sequência 00111000 e a letra ‘e’ (minúsculo) corresponde à sequência 01100101. Este código binário, conhecido como código de máquina, é o único que os processadores são capazes de compreender. Porém, dar instruções ao

² Grosso modo, máquinas clássicas se opõem às máquinas quânticas. Repare que as máquinas clássicas funcionam com a transmissão de *bits*, que podem variar entre os estados desligado (0) ou ligado (1). As máquinas quânticas funcionam com os chamados *qubits*, ou *bits* quânticos, na qual a informação, além dos estados binários, pode estar em uma sobreposição de 0 e 1 (vetores). Isso terá efeitos na quantidade de informação transmitida por vez e na forma de transmissão e leitura de informação pelos computadores quânticos.

computador diretamente em binários é impraticável. Por isso, surgiram as linguagens de programação, doravante LP, visando agilizar a tarefa do desenvolvedor.

As LPs são classificadas de acordo com seu nível de abstração. As linguagens de mais baixo nível têm um funcionamento mais próximo do conjunto de instruções suportado pelos processadores das máquinas, o que exige uma maior curva de aprendizagem. As linguagens de mais alto nível são mais próximas da linguagem humana utilizando funções e relações sintáticas e semânticas de mais fácil memorização e aprendizagem (Figura 2).

FIGURA 2 – Linguagens de programação e nível de abstração



Nota: O primeiro passo de um curso de programação é escrever um código que faça a máquina exibir a frase “*Hello World*” no monitor. A Figura A corresponde apenas às letras do texto em binários; A Figura B utiliza um Assembly para realizar esta exibição; a Figura C apresenta o mesmo comando na linguagem C; A Figura D, executa a exibição do texto em Python. Repare que a facilidade de escrita do código aumenta de A até D. A seta maior exemplifica os conceitos de linguagens de baixo nível de abstração (mais próximo dos binários) e as de alto nível (mais próximos da linguagem humana). A forma como devemos escrever e organizar as funções e argumentos é chamada de sintaxe.

Como exemplo de linguagem de baixo nível é necessário citar o Assembly. O Assembly não é exatamente uma linguagem específica, mas o nome que se dá a uma linguagem única de cada processador,

contendo uma forma legível e memorizável do conjunto de instruções que a máquina pode realizar. Por esta razão o Assembly é também conhecido por ‘linguagem de montagem’. Programadores mais avançados podem utilizar o Assembly para dar instruções para a máquina ou mesmo para desenvolver aplicações. O porém do desenvolvimento em Assembly é que eles são ininteligíveis as outras máquinas e seus códigos só podem ser executados pelo modelo de processador para o qual foram escritos. Além disso, a programação em Assembly ainda é uma tarefa para experts na área.

A programação só viria a se popularizar com o desenvolvimento de linguagens de alto nível de abstração. Por outro lado, assim como em linguagem natural, quando dois interlocutores não são fluentes em uma língua em comum, é necessário um tradutor/intérprete para que a comunicação se estabeleça. Isso gera um custo de processamento e, por consequência, um aumento no tempo de resposta da máquina.

Uma linguagem é um software que nos permite criar, de forma lógica, uma sequência de passos/funções chamada ‘algoritmo’.³ Este algoritmo será lido e executado pela máquina, esta, que só entende binários. Para que estas instruções sejam compreendidas pelo hardware, a linguagem precisa ser ‘**compilada**’, traduzida para a linguagem de máquina (ex. C e C++), ou ser ‘**interpretada**’, transformando suas linhas em um código binário (*byte code*) que será interpretado por uma máquina virtual (ex. Java e Python).

Uma das linguagens mais utilizadas hoje é o ‘C’. Por ter sido criado com o objetivo de desenvolvimento de sistemas operacionais, softwares que precisam tirar todo o potencial das máquinas, o C é considerado a língua franca da programação, assim como o inglês entre as línguas naturais. As fabricantes de hardware, além do Assembly, geralmente escrevem também um código que mapeia as funções de seus dispositivos diretamente para o C, facilitando o trabalho dos desenvolvedores. Apesar da praticidade introduzida por esta linguagem, a programação ainda era relativamente restrita. Além disso, muitos usuários têm necessidades bastante específicas que poderiam ser realizadas de forma mais simples e em uma lógica de programação diferente da utilizada pelo C. E assim

³ Algoritmo pode ser, grosso modo, definido como uma sequência ordenada de passos que levam a um determinado resultado, logo, se trata de uma sequência finita. Um algoritmo não precisa necessariamente ser algo matemático ou computacional.

surgiram diversas linguagens de mais alto nível, das quais podemos citar o Python, o R, o Matlab e o Java. A vantagem destas linguagens é que todas permitem que executemos cada passo do código ao longo da programação para verificar se ele irá funcionar corretamente na prática. Isso facilita (i) a identificação e a localização de erros, que normalmente são indicados pelo próprio console (tela de escrita do código) durante a programação, e (ii) a aprendizagem, devido ao seu *feedback* imediato e identificação dos argumentos sintáticos através de diferentes cores, como podemos observar no exemplo do Python, na Figura 2.

O Java é bastante utilizado pois, além de ser gratuito, teve uma grande ação de marketing realizada pela sua desenvolvedora Sun, hoje pertencente à Oracle. O mercado passou a exigir conhecimentos de Java para contratar programadores, o que levou as universidades a lecioná-la nos cursos ligados à tecnologia. Embora tenha perdido força recentemente, o Java é a linguagem de programação adotada pela Google para desenvolvimento de aplicativos para Android. O Java é interpretado e traduz o código para uma máquina virtual (*Java Virtual Machine*, JVM). Esta máquina virtual é uma espécie de emulador que simula uma máquina específica em qualquer computador, evitando que o código tenha que ser recompilado.

O Python funciona de forma semelhante ao JAVA, utilizando uma máquina virtual que é instalada em qualquer computador e permite que ele seja executado em qualquer máquina, independente do sistema operacional. Isso favorece a portabilidade de seus códigos e o torna uma das linguagens preferidas de quem realiza experimentação.

O Matlab é um software proprietário com base em C e em Java. Ele foi criado com o objetivo facilitar a programação baseada em matrizes de dados. Embora os códigos escritos em Matlab rodem diretamente dentro do software, ele possui um compilador que nos permite tornar estes softwares independentes, capazes de rodar fora de sua interface. Uma opção gratuita ao Matlab é o GNU Octave. Sua sintaxe é bastante semelhante à do Matlab, o que facilita a migração. Outra opção com sintaxe semelhante é o Julia (BEZANSON *et al.*, 2014) que, em algumas tarefas, apresenta um excelente desempenho.

Já o R é bastante conhecido de qualquer linguista e outros pesquisadores que trabalham com estatística, análise ou mineração de dados para a criação de corpus. Trata-se de um software livre desenvolvido para facilitar o trabalho com dados numéricos e estatísticos, sendo

amplamente utilizado em ciências cognitivas. Psicolinguistas costumam utilizar o R para análise de dados devido à sua lógica de programação e aos diversos códigos distribuídos gratuitamente para tal, mas nada impede que ele seja utilizado para desenho e aplicação de experimentos.

Em resumo, observamos até aqui que o processador possui uma arquitetura que recebe um determinado tipo de informação para executar um algoritmo. Essa informação pode ser elaborada através de uma linguagem de programação que, além de facilitar a tarefa do programador, também pode ser traduzida para a máquina através de compiladores e interpretadores, em troca de uma determinada perda no desempenho do algoritmo. Agora que temos um conhecimento básico sobre a comunicação entre hardware e software, seguimos com nossa discussão sobre os softwares desenvolvidos para a criação e apresentação de estímulos em ciências cognitivas.

4. Softwares para experimentação em ciências cognitivas

Uma das maiores dificuldades de um aluno que escolhe trabalhar com o método experimental é aprender como controlar devidamente a estimulação e a coleta dos dados. Além disso, reparo que, no Brasil, boa parte dos experimentos que são realizados em software proprietários de mais de mil dólares poderiam ser portados com facilidade para software livre. Por este motivo, proponho uma mudança na relação dos psicolinguistas com os softwares especializados em experimentação.

Neste caminho, nos deparamos com quatro problemas básicos que, embora sem uma ordem específica, serão considerados e discutidos ao longo desta seção:

- (i) **Escolha:** a grande oferta de softwares de estimulação existentes hoje no mercado;
- (ii) **Familiaridade** com a tarefa: a falta de familiaridade com conceitos mais básicos na relação entre hardware e software e com linguagem de programação;
- (iii) **Aprendizagem:** a curva de aprendizagem em algumas plataformas de experimentação; e
- (iv) **Portabilidade:** a diferença entre os sistemas operacionais que, frequentemente, limita a escolha de softwares.

No que diz respeito à escolha, há algumas décadas, as opções disponíveis para elaborar testes psicométricos eram escassas, obrigando os iniciantes a utilizar os recursos já disponíveis em seu laboratório. Isso facilitava o fator escolha mas afetava o fator aprendizagem, visto que precisamos nos acostumar com os softwares que eram disponibilizados. Hoje, vivenciamos um crescimento nas opções de softwares especializados em experimentação, nos trazendo cada vez mais opções de escolha.

Hoje é comum os laboratórios americanos e europeus disponibilizarem de três a quatro opções para facilitar o trabalho de integrantes e visitantes que, porventura, tenham experiência anterior em algum deles. Ao solicitar um posdoc, é também comum os laboratórios exigirem experiência em um ou dois softwares específicos, correspondentes àqueles nos quais seus integrantes mais trabalham. Isso traz uma uniformidade na forma como o laboratório desenvolve suas pesquisas. Por outro lado, o que deveria ser uma facilidade, em certos casos, se torna um problema.

Hoje existem dezenas de softwares que podem ser utilizados para a elaboração de experimentos. Alguns destes softwares nos oferecem desde a liberdade criativa das linguagens *Turing complete*⁴ como C, Presentation, Java, R, Python e Matlab e suas toolboxes (caixa de ferramentas). Outros nos dão a facilidade de aprendizagem e de uso em detrimento da liberdade de criação nos softwares proprietários com interface gráfica (GUI, de *Graphic User Interface*), como o E-Prime, o Paradigm e o SuperLab. Outros ainda chegam a combinar a praticidade da GUI, a liberdade da programação e, também, a portabilidade entre sistemas operacionais, como o PsychoPy e o OpenSesame, ambos em software livre.

O maior questionamento na escolha de um software cai em cima dos iniciantes. Ainda inexperientes e tendo que dividir suas atenções entre graduação, iniciação científica e pré-projetos para obter bolsas e para ingressar em um mestrado, eles poderão apresentar ainda mais

⁴ Mais tecnicamente, existem diversos tipos de linguagens. Segundo Alan Turing, uma linguagem de programação deve possuir (i) uma forma de repetição ou de salto condicional e (ii) um fim, possibilitando a geração e a leitura de um resultado do algoritmo programado. Ao atender estas condições, a linguagem é chamada de *Turing Complete*. Linguagens *turing complete* nos permitem programar tudo o que nossa habilidade nos permitir.

dificuldades tanto na escolha, quanto na curva de aprendizagem e na habituação a um software durante seus primeiros testes. Neste momento, mesmo aqueles que consideram ter um bom nível de conhecimento, muito provavelmente optarão pela opção mais prática, independente de ela ser, de fato, mais prática para ele, ou de ser a melhor opção para o tipo de teste que irá aplicar.

Como vimos anteriormente, as linguagens de programação nos permitem criar tudo o que nossa habilidade como programadores permitir. Desta forma, é perfeitamente possível utilizá-las para criar qualquer tipo de experimento, desde os mais simples e recorrentes até experimentos mais complexos e com métodos completamente inéditos. Ainda assim, escrever todos os comandos necessários para a comunicação entre software e hardware, além de reescrever funções corriqueiras podem dificultar e alongar este trabalho. Para facilitar a tarefa, diversos grupos de pesquisadores com habilidades em programação desenvolveram softwares que facilitam a vida do programador-experimentador.

Nas próximas subseções apresento e discuto algumas destas opções. Comparações técnicas entre cada uma delas, como exatidão temporal na coleta de dados ou velocidade de processamento, porém, estão fora do escopo deste trabalho, especialmente porque estas medições podem mudar de acordo com o hardware utilizado. Para garantir que os tempos apresentados não foram alterados por conta de problemas de sistema e/ou de hardware é necessário a utilização de medidores externos (PLANT *et al.*, 2004; PLANT; TURNER, 2009). Caso estas comparações sejam interessantes para você, sugiro iniciar pela bateria de testes realizada por Garaizar *et al.* (2014), comparando os softwares E-Prime 2, PsychoPy e DMDX, ou o artigo de De Leeuw & Motz (2015) que comparam os tempos de resposta do PsychToolbox 3 com o do jsPsych rodando em navegadores e advogando pela viabilidade de experimentos cronométricos via internet.

4.1 Opções de software experimental #1: As *toolboxes*

Uma das opções que nos auxiliam no desenvolvimento dos experimentos são as *toolboxes*. Uma toolbox consiste em uma ‘caixa de ferramentas’, contendo um conjunto de funções previamente escritas para determinados fins. Com estas ferramentas, não precisamos realizar toda a comunicação entre hardware e software, bastando utilizar estas funções e definir seus parâmetros em nosso código, o que facilita e

automatiza as tarefas mais comuns, no nosso caso, tarefas de apresentação de estímulos e de coleta de dados. O quadro abaixo, nos mostra algumas das toolboxes utilizadas para experimentos em ciências cognitivas, o que inclui a psicolinguística e a neurociência da linguagem:

QUADRO 1 – Algumas opções de toolboxes utilizadas para experimentação em ciências cognitivas

Para Python	
• <i>ExPyrimint</i>	Krause & Lindermann (2014)
• <i>PyGame</i>	Shinners (2011)
• <i>Vision Egg</i>	Straw (2008)
Para C	
• <i>PsyToolKit</i>	Stoet (2010)
Para Java	
• <i>PsychJava</i>	www.psychjava.com ⁵
Para Matlab	
• <i>Psychtoolbox 3</i>	Kleiner <i>et al.</i> (2007)
Para JavaScript (Experimentos via Web)	
• <i>jsPsych</i>	De Leeuw (2014)
• <i>JATOS</i>	Lange; Kühn; Filevich (2015)

O *ExPyrimint* (KRAUSE; LINDERMANN, 2014) foi elaborado para ser uma plataforma universal de experimentação. Primeiramente por ter funções para experimentos comportamentais e neurofisiológicos, segundo por ser multiplataforma, graças ao Python. Sua ideia é ter uma lógica estruturada de forma a facilitar a transposição do desenho experimental para o seu código. Uma de suas vantagens é a possibilidade de rodar uma versão específica para *tablets* e *smartphones* android. Em seu site é possível encontrar tutoriais e uma série de códigos modelo para serem estudados e utilizados na programação de seu teste (ver Anexo).

⁵ O site do PsychJava está fora do ar há algum tempo. Não consegui informações sobre os motivos da queda do site. Como ele já estava incorporado em outros softwares, acredito que o projeto esteja parado.

Os desenvolvedores do *Vision EGG* (STRAW, 2008) buscavam uma forma mais simples de utilizar a linguagem de programação multiplataforma Python para o processamento gráfico, especialmente em 3D. Eles então aproveitam o conjunto de bibliotecas de funções (API⁶) gráficas conhecido como *OpenGL* e desenvolvem o *VisionEGG* como uma interface entre os dois, visando a experimentação com estímulos visuais.

O *PyGame* (SHINNERS, 2011) tem uma proposta diferente. Originalmente criado para desenvolvimento de jogos, sua ideia é que o seu próprio código seja responsável por tarefas que demandam maior poder de processamento como a renderização de imagens e processamento de áudio. Estas tarefas serão abstraídas de forma diferente do código escrito pelos desenvolvedores, garantindo sempre o melhor desempenho possível para o jogo. Estas características o tornaram uma boa ferramenta para desenvolvimento de experimentação, sendo considerado uma boa plataforma para estimulação visual e auditiva, que demandam mais processamento da máquina.

Uma das toolboxes mais utilizadas atualmente é a *Psychtoolbox 3*, ou PTB-3, (KLEINER *et al.*, 2007), desenvolvido para Matlab e GNU Octave. A proposta do PTB-3 é fornecer funções que façam interface entre o Matlab e o hardware para fins de experimentação. Isso permite um maior controle e confiabilidade cronométrica das estimulações visuais e auditivas apesar de ser uma linguagem mais distante do hardware (de alto nível de abstração). Estas características a tornam uma excelente ferramenta para programadores iniciantes. O PTB também tem interface com a API gráfica *OpenGL*, além de ter funções escritas por fabricantes de hardware como a *EyeLink Toolbox*, fornecida pela SR Research para desenvolvimento de testes em seus equipamentos de rastreamento ocular.

Embora o Matlab tenha versões nas diferentes plataformas, é bastante provável que alguns códigos precisem ser ligeiramente modificados para se tornar compatíveis com um novo sistema operacional

⁶ O termo API (*Application Programming Interface*) se refere a um conjunto de algoritmos criados pelo desenvolvedor de um software para permitir que outros softwares, ou um código criado pelo próprio usuário, utilizem ou modifiquem algumas funções ocultas da aplicação em questão.

como, por exemplo, o mapeamento do teclado.⁷ Apesar disso, devido à grande adoção, o PTB-3 é constantemente atualizado para corrigir bugs, para aumentar suas funções e para melhorar seu desempenho e compatibilidade. Mais além, a toolbox contém funções do *PsychJava* que ainda não possui distribuição pública.

4.2 Opções de software experimental #3: Opções de experimento via Web

Caso você tenha a necessidade de realizar um experimento em massa ou, por alguma outra razão, a experimentação via web seja interessante para você, existe a possibilidade de utilizar linguagens específicas para desenvolvimento de páginas web, como o HTML5, o CSS e o JavaScript. Outra opção é o Flash que, por diversas razões técnicas e, muito provavelmente também, estratégicas e de mercado, se tornou indesejável pelo mercado.

Estas linguagens também possuem suas toolboxes, facilitando bastante a tarefa de elaboração de testes psicofísicos na web. Uma toolbox bem recente é o *jsPsych* (DE LEEUW, 2014), para *JavaScript*. O *jsPsych* disponibiliza alguns modelos que podem ser reutilizados para outros tipos de testes, o que facilita bastante o seu uso. Aqueles que já trabalharam com JavaScript, CSS e HTML5, linguagens desenvolvidas para criação de páginas web, provavelmente terão facilidade em desenvolver experimentos com esta toolbox. Outra toolbox com o mesmo objetivo e utilizando a mesma linguagem é o *JATOS* (*Just Another Tool for Online Studies*), de Lange, Kühn & Filevich (2015).

Também temos opções em outras linguagens. Desenvolvido para C, o *PsyToolkit* foi criado por Gijsbert Stoet para criação e aplicação de experimentos comportamentais. A ideia deste *toolkit* é ser uma linguagem de nível de abstração maior que o do C, tendo um compilador duplo que transmuta o código para C durante a programação para, em seguida, o compilador do C transformá-lo em linguagem de máquina. Desde sua versão 1.4 é possível interpretá-lo na máquina virtual Java (JVM). Este *toolkit* também possui uma interface web que permite criar e rodar experimentos via web, além de uma interface gráfica para criação e aplicação de questionários online (*PsyQuest*). Em seu site é possível

⁷ A mudança no código das teclas no PsychToolbox 3 pode ser observada nos códigos anexos de Sampaio (2015). Dos 4 experimentos desenvolvidos com a *toolbox*, 2 foram aplicados em PCs e 2 em Mac.

encontrar diversos tutoriais e modelos dos paradigmas mais populares em experimentação (ver Anexo).

Embora alguns testes possam ser facilmente portados para plataformas Web, ainda sou bastante cético no que diz respeito ao controle da estimulação. Alguns estímulos visuais e auditivos podem variar bastante de acordo com o monitor, caixas de som, fones e hardware utilizados. Máquinas e navegadores diferentes podem apresentar o experimento e coletar os dados cronométricos de forma diferente. Além disso, não temos um bom controle sobre as condições ambientes e sobre quem são os participantes do teste em casos em que estas informações sejam relevantes na interpretação dos dados.

No que diz respeito aos tempos de reação, De Leeuw & Moritz (2015) realizaram uma bateria de testes comparando o desempenho do jsPsych com o PsychToolbox 3 e advogam a favor da utilização de JavaScript inclusive para testes cronométricos. Já Reimers & Steward (2014) comparam testes em JavaScript em Flash. Os autores argumentam que ambos podem ser ferramentas úteis para experimentação psicofísica. Nos últimos anos, porém, o Flash vem sendo excluído de ambiente web, o que me faz acreditar que, mesmo que ainda seja uma ferramenta útil, é possível que, em breve, testes escritos em Flash deixem de ser viáveis. De qualquer forma, o Flash gera arquivos bastante pesados em relação aos outros softwares, o que pode comprometer o desempenho em máquinas mais antigas e menos potentes.

Outra opção interessante para realizar experimentos na web e em massa é o desenvolvimento de testes para tablets, que vêm se tornando uma ferramenta cada vez mais explorada. Experimentos para *tablets* podem ser desenvolvidos diretamente em Java (android) ou Swift (iPad), além de poderem ser desenvolvidos em outros softwares livres ou proprietários, como veremos nas próximas seções. Para iPad, ainda existe a opção de desenvolvê-lo no *PsyPad*, criado e mantido por Andrew Turpin (TURPIN; LAWSON; MCKENDRICK, 2014).

4.3 Opções de software experimental #3: Linguagens específicas para experimentos cognitivos

As toolboxes facilitaram muito o trabalho de desenvolvimento de experimentos cognitivos em diversas linguagens de programação. Porém, se os programadores criam linguagens próprias para facilitar suas próprias

tarefas, como o R e o Matlab, os pesquisadores em ciências cognitivas também criaram linguagens que facilitam a apresentação de estímulos.

Este é o caso do PEBL (*Psychology Experiment Building Language*; MUELLER; PIPER, 2014), baseado em C++, gratuito, desenhado especificamente para a elaboração de experimentos com estímulos de texto, imagens, áudios e vídeos. Esta linguagem está disponível para Windows e MacOS e sua utilização consiste na criação e na edição de arquivos de texto modelo que contém os códigos necessários para que o *parser* da linguagem de programação apresente os estímulos e colete os dados indicados pelo programador (ver Anexo). Outra opção gratuita é o DMDX (FORSTER; FORSTER, 2003), bastante utilizado para experimentos visuais.

Outros softwares deste tipo foram desenvolvidos por empresas e são, portanto, pagas. Um dos softwares proprietários mais utilizados nas últimas décadas é o *Presentation*, elaborado pela *Neurobehavioral Systems*. O *Presentation* contém duas linguagens proprietárias, a SDL (*Scenario Description Language*), e a PCL (*Program Control Language*), baseadas em C e em Basic, ambas utilizadas para elaborar os estímulos visuais, trials e o roteiro da estimulação. Atualmente, o *Presentation* conta com um módulo que permite a programação em Python.

4.4. Opções de software experimental #4: *Graphic User Interface* (GUI)

Apesar das facilidades introduzidas pelas linguagens de programação, pelas toolboxes e também pelas linguagens mais direcionadas à experimentação, tudo isso ainda envolve o ato de programar que, para alguns, ainda é considerado uma tarefa de especialistas. Iniciantes e profissionais mais experientes em ciências cognitivas, que não tiveram formação em lógica de programação, possuem uma enorme resistência à necessidade de programar seus experimentos. Para eles, foram elaborados alguns softwares que oferecem uma interface gráfica (GUI), que torna o processo de desenvolvimento mais visual, diminuem a necessidade de habilidades de programação e, assim, diminuem também a curva de aprendizagem necessária para criar seus primeiros testes.

Um dos softwares GUI mais famosos é o *Psyscope* (COHEN *et al.*, 1993), bastante utilizado para experimentação em linguagem. O *Psyscope* possui uma interface gráfica, com objetos *drag-and-drop*, que permitem visualizar e organizar experimentos em uma lógica visual de

diagrama arbóreo. As linhas indicam as relações entre funções, listas e objetos do experimento, cada um com uma gama de opções internas que nos dão enorme liberdade de configuração e personalização do nosso teste.

A versão atual do Psyscope conta com suporte aos aparelhos de rastreamento ocular da Tobii. Embora rode nativamente em processadores Intel,⁸ o Psyscope ainda é exclusivo do MacOS, o que se configura em uma desvantagem, especialmente no que diz respeito ao preço do equipamento. Apesar disso, ele tem como vantagem o fato de ser gratuito e de que é possível dominá-lo em questão de dias. Uma nova versão, ainda em fase beta, possui um editor de código interpretado. Esta mudança deve permitir a identificação de erros e a alteração de determinadas funções de forma muito mais simplificada, via código. Aos experimentadores que desejarem testar a nova versão, basta entrar em contato com Luca Bonatti, um dos desenvolvedores e responsáveis pelo fórum (ver Anexo).

Em ambiente Windows, um dos softwares mais próximos ao Psyscope é o E-Prime, de código proprietário. O E-Prime também possui uma interface *drag-and-drop* na qual é possível organizar e visualizar o experimento, porém, sua lógica simula uma linha de tempo, na qual listas e funções se sucedem. Sua versão 3.0 foi lançada em dezembro de 2016 com a possibilidade de desenhar experimentos para *tablets*. Devido ao lançamento recente, os comentários sobre o E-Prime neste artigo se referem à versão 2.

Outro software semelhante é o *Paradigm*.⁹ Com lógica semelhante e baseado em Python. Ele tem a vantagem de possibilitar a criação de experimentos que podem ser salvos em *DropBox* para serem apresentados em iPads. Tanto o E-Prime quanto o *Paradigm* contam com suporte das fabricantes. Os preços, porém, são uma grande desvantagem.

⁸ A maior razão da incompatibilidade de diversos softwares entre plataformas Mac e PC era a utilização de processadores diferentes. Hoje todas as máquinas da Apple utilizam processadores Intel, o que permite, por exemplo, que o Windows seja instalado em um Mac, a existência dos diversos emuladores de Windows no Mac e, também, dos chamados Hackintoshs, que consistem na instalação do MacOS em PCs. Por esta razão, hoje o Psyscope poderia ser portado para o Windows, o que ainda não foi feito pelos desenvolvedores.

⁹ No início de 2016 o *Paradigm* teve suas vendas interrompidas por conta do falecimento de seu único desenvolvedor, Bruno Tagliaferri. A empresa foi comprada por Josh Pritchard no final do mesmo ano, retornando as vendas e o suporte.

Além destes, existem softwares que são desenvolvidos pelas próprias fabricantes dos equipamentos, de forma a certificar a comunicação eficiente entre software e o seu hardware. Este é o caso dos rastreadores oculares. Para citar apenas as maiores fabricantes, o equipamento da Tobii conta com o software Tobii Studio, uma ferramenta que segue a lógica da linha do tempo para organizar os estímulos visuais. O equipamento da SMI conta com toda uma suíte de aplicativos para desenhar, aplicar e analisar os dados. Já os rastreadores da EyeLink contam com o software *Experiment Builder*, além de toolboxes adicionais para o Psychtoolbox, do Matlab e para Python.

A grande vantagem dos softwares com interface gráfica está na curva de aprendizagem. Geralmente um iniciante consegue aprender a usar e montar seu experimento em poucos dias. Porém, uma desvantagem é o fato de serem muito focados no seu objetivo principal: realizar experimentação. Desta forma, embora criem diversos tipos de algoritmos bastante poderosos e avançados, eles são apenas softwares de experimentação, não nos permitindo ir muito além das funções já previstas pelos seus idealizadores.¹⁰

Alguns softwares que driblam esta questão vêm surgindo no mercado, oferecendo uma interface gráfica que facilita a visualização da sequência de algoritmos, mas, ao mesmo tempo, por serem baseados em linguagens de alto nível, conseguem mesclar sua interface gráfica com o potencial da programação. Felizmente, as duas opções que conheço são gratuitas e multiplataforma: o *PsychoPy* e o *Open Sesame*, ambos baseados em Python.

¹⁰ Excluindo o Psyscope, tecnicamente E-Prime e Paradigm podem ser expandidos através das ferramentas chamadas “*InLine*”. Esta ferramenta permite inserir pedaços de programação em outra linguagem dentro do código gerado pelo software GUI. Os comandos *InLine* são a forma como podemos acessar algumas funções escondidas dos softwares, tendo apenas o objetivo de extensão das possibilidades oferecidas na interface gráfica. Desta forma, a linguagem dos *InLine* não é utilizada para criar um código completamente novo com funções que já não foram, de alguma forma, inseridas pelos desenvolvedores do software. Quando uma linguagem é inserida dentro de outros softwares com esta finalidade, elas são conhecidas como *Linguagens de Script* e criam *scripts*, diferente do código que o software cria ao final do processo de desenvolvimento e que contém estes scripts.

O PsychoPy¹¹ (PEIRCE, 2007, 2009) possui uma interface gráfica que permite elaborar e visualizar a organização de uma grande parte do seu experimento. Grosso modo, ele possui duas janelas de linhas do tempo, uma do experimento como um todo e outra de cada estímulo a ser apresentado. Este software roda em uma base (*backend*) que faz interface entre o Python e o OpenGL (pyglet).

A facilidade de uso e o fato de rodar em Python trouxe aos usuários de PsychoPy o sonho de vê-lo rodando em um RaspberryPi, uma espécie de minicomputador desenvolvido pela Fundação RaspberryPi no Reino Unido (Figura 3). Estes computadores são extremamente baratos, custando menos de 40 dólares sua versão mais potente hoje (versão 3 Model B) e menos de 20 dólares em sua versão mais simples (versão Zero). Devido ao seu preço, estes computadores vêm se tornando popular em todo tipo de projeto que envolva recursos computacionais. Porém, devido a incompatibilidades entre software (pyglet) e hardware do RaspberryPi, o PsychoPy era incompatível com o RaspberryPi. Este panorama pode mudar em breve. Felizmente, foram lançados em no ano passado os primeiros drivers experimentais do OpenGL para a plataforma, possibilitando o uso do PsychoPy nestes pequenos computadores. Segundo testes realizados por Mark Scase e publicados no fórum do PsychoPy¹² em fevereiro de 2016, ainda é inviável aplicar experimentos. Mas ainda assim é possível criar códigos no RaspberryPi e aplicá-lo em máquinas com drivers mais funcionais.

¹¹ É comum que categorizem o PsychoPy como uma *toolbox* devido a algumas características. Não discordo, porém, o fato de ele possuir uma GUI faz com que ele tenha mais interessados entre os leigos do que as *toolboxes* tradicionais e, por isso, preferi categorizá-lo entre os softwares GUI.

¹² “PsychoPy on RaspberryPi”: <https://groups.google.com/forum/#!topic/psychopy-users/1mPwJqDVy1c>

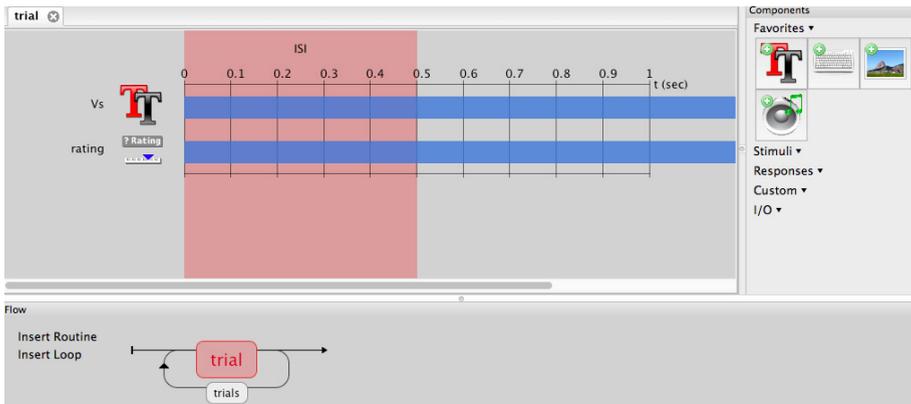
FIGURA 3 – Um RaspberryPi 3 Model B, em uma case de proteção (foto autoral).



Apesar da facilidade e de uma rápida curva de aprendizagem,¹³ ainda me parece mais simples configurar as variáveis de algumas funções diretamente no código do PsychoPy. Outras, podem realmente não estar disponíveis na interface gráfica visto que o software privilegia uma interface simples com as funções mais comuns em experimentação psicofísica. Por exemplo, ele não possui um editor de tabelas em sua interface como o E-Prime e o Psyscope, por mais que estes editores sejam bastante limitados. Desta forma, é necessário organizar nossas tabelas em um software externo como o Excel. Geralmente isso é um procedimento padrão para alguns programadores e simples para quem está iniciando na área, não trazendo nenhuma dificuldade extra. Ainda assim, entre os utilizadores da interface gráfica, esta ausência é normalmente apontada como um de seus pontos fracos.

¹³ Excelente tutorial do PsychoPy em Português gravado pela Prof. Mahayana Godoy (UFRN): <www.youtube.com/watch?v=W8cpnARvtNw>.

FIGURA 4 – Captura de tela da interface gráfica do PsychoPy



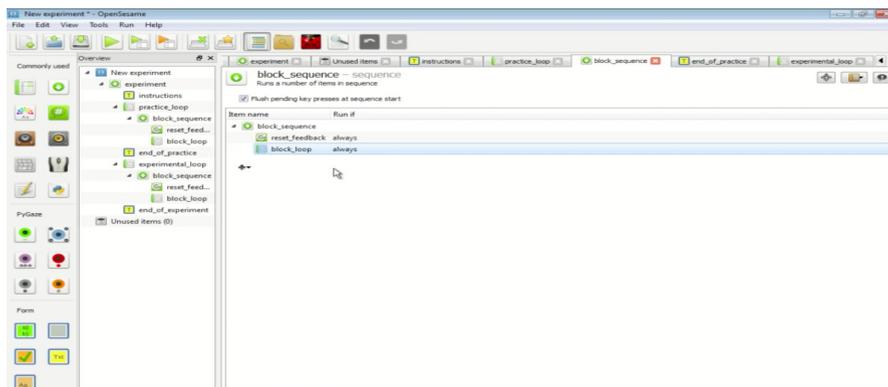
Nota: Além da Interface GUI ele conta com um console para programação em Python.

O Open Sesame¹⁴ (MATHÔT *et al.*, 2012), embora tenha uma interface mais completa que a do PsychoPy, ainda requer algum conhecimento da relação hardware-software para que seja corretamente utilizado. Um exemplo é a decisão sobre o tempo dos estímulos visuais em comparação com a taxa de atualização de telas do monitor utilizado, como discutiremos com mais detalhes na próxima seção. Outra questão é a escolha da melhor base (*backend*) para ser utilizada pelo OpenSesame para o seu experimento, de acordo com o tipo de teste e de hardware [pyglet, pygame, xpyriment ou droid¹⁵]. Caso o pyglet não seja necessário, o Open Sesame pode se tornar uma boa ferramenta para ser utilizada nos RaspberryPi. Este software pode ser considerado também uma opção gratuita ao E-Prime visto que sua interface possui alguma semelhança com a do software proprietário.

¹⁴ Na prática, o Python funciona como uma Linguagem de Script no Open Sesame, aparentemente de forma distinta do que acontece no PsychoPy. Ainda assim, o Python é mais fundamental no Open Sesame do que as linguagens de script no E-Prime ou no Paradigm e, por este motivo, o categorizamos entre as *toolboxes* e os softwares GUI.

¹⁵ Para experimentos em tablets Android.

FIGURA 5 – Captura de tela da Interface gráfica do Open Sesame



5. Cuidados na relação software-hardware que influenciam a percepção

A seção anterior traz uma gama de alternativas de software para elaborar a estimulação. Por outro lado, nossos cuidados não podem se resumir a uma boa escolha de software. Assim como Chomsky propõe a diferença entre a competência e o desempenho, separando o que sabemos do que de fato fazemos em linguagem, podemos transpor a dicotomia para a relação software-hardware. O software nos permite enviar um comando para que a máquina faça uma determinada tarefa, mas será que o hardware é capaz de realizá-la?

5.1 Cronometria e estimulação visual: o caso do movimento aparente

Após os esforços de Helmholtz na Psicologia Fisiológica e de sua recuperação por Donders e Cattell na Psicologia Experimental, nossos métodos de coleta e de análise de tempos de resposta (Cronometria Mental) são reconhecidamente uma ferramenta de análise e de medida dos processos cognitivos. A psicolinguística se utiliza frequentemente de protocolos cronométricos em modalidades visual e auditiva como nos experimentos de decisão lexical, de priming, de leitura automonitorada, em testes de percepção entre outros. Muitos testes, porém, dependem de exatidão temporal na escala dos milissegundos e, para isso, é necessário ter uma noção do funcionamento de nossos equipamentos, como o monitor.

Antes de entrar nos detalhes do funcionamento dos monitores, precisamos entender duas ilusões visuais que foram de extrema importância na história do seu desenvolvimento. A primeira delas é o *Phi Phenomenon* (WERTHEIMER, 1912), que ocorre quando dispomos diversas lâmpadas uma ao lado da outra e as ligamos e desligamos de forma sucessiva. Esta ação impede que nossa mente perceba o desligar e ligar das lâmpadas, criando uma ilusão de que a luz se move de uma lâmpada para a outra.

A segunda ilusão é a *Beta Movement*, descrita por (KENKEL, 1913). Se apresentamos uma sequência de imagens ligeiramente semelhantes – como um boneco em diferentes posições – em uma determinada velocidade, nossa mente não consegue concebê-las como imagens estáticas, mas como uma mesma imagem cuja cena está em movimento. Estes dois fenômenos nos causam a ilusão conhecida como Movimento Aparente.

Estas ilusões são as responsáveis pela nossa capacidade de nos entreter com vídeo games, animações e filmes. Duas questões foram postas para as técnicas de apresentação de imagens com movimento aparente: (i) criar um material com maior número de imagens para resultar em uma melhor experiência ou (ii) criar um material que mantenha a experiência aceitável da forma mais barata possível?

Nos primeiros filmes mudos, as imagens eram apresentadas em uma sequência de frames registrados em películas de celuloide numa taxa que variava entre 14 e 26 quadros por segundo (fps, *frames per second*), que foi o suficiente para dar a ilusão de movimento. Por outro lado, este movimento era normalmente considerado irregular, dando a sensação de que as imagens tinham pequenos saltos no tempo (*skipping*). Desta forma, podemos dizer que o *threshold*¹⁶ para o *beta movement* é de aproximadamente 15fps. Para resolver este problema, os filmes passaram a ser gravados e apresentados em uma taxa mais alta, variando entre 18 e 23fps, melhorando consideravelmente a experiência dos vídeos. Mais tarde, essa taxa passou para 24fps fixos, visto que esta é a taxa mínima

¹⁶ Podemos traduzir *threshold* pelo termo “limiar”. O limiar seria um ‘nível’ a partir do qual podemos observar uma mudança na percepção. Neste caso, até 14fps o sistema visual humano consegue distinguir as imagens de um filme. A partir de 15fps aproximadamente, esta percepção passará a ser de movimento, embora uma melhora na fluidez do movimento possa ainda ser observada com o aumento da frequência.

para que os vídeos pudessem ser corretamente sincronizados com o som (READ; MEYER, 2000).

5.2 Por que usar monitores CRT?

A televisão foi inventada na década de 50 e levou as imagens de filmes e programas televisivos para dentro das casas. Estes aparelhos eram enormes e pesados devido a sua tecnologia. Existem elementos que emitem radiações através da absorção de uma fonte de energia. Este é o caso do fósforo, que é utilizado tanto em objetos fluorescentes, que emite radiação visível enquanto absorve radiações de outras fontes, quanto nos fosforescentes, que continuam a emitir radiação visível algum tempo após a absorção. As telas das TVs são fosforescentes e absorvem os elétrons emitidos por um grande tubo de raios catódicos (*Cathode Ray Tube*, ou CRT), responsáveis pelo tamanho e peso destes aparelhos.

Os monitores mais antigos seguem a mesma tecnologia. Nos monitores CRT, cada quadro (imagem estática) é construído pixel por pixel de forma sequencial, partindo do primeiro ponto no canto superior esquerdo até o último no canto inferior direito da tela, tudo isso em poucos milissegundos. Neste momento, o computador recebe um sinal do monitor indicando que o quadro atual foi finalizado e começa a construção do próximo quadro. Este sinal é chamado de *retrace signal* (sinal de retorno; COHEN; PROVOST, 1994). Para não percebermos a mudança entre quadros, a tela pisca durante 1,5ms enquanto os raios que iluminam cada pixel da tela retornam ao canto superior esquerdo para iniciar a montagem do próximo quadro (PEIRCE, 2009). A frequência na qual um monitor consegue trocar de um quadro para o outro ficou conhecido como *refreshrate* (taxa de atualização). Este termo, em parte, substituiu o *fps* nas descrições de manuais.

Uma curiosidade dos aparelhos de TV é que o *refreshrate* era definido de acordo com a alternância de corrente elétrica local. Nos EUA, essa frequência é de 60Hz enquanto o fornecimento de energia na Europa funcionava a 50Hz. O *refreshrate* indica o quão rápido um aparelho consegue atualizar a imagem a cada segundo. Assim, os aparelhos de televisão na Europa poderiam apresentar uma imagem diferente a cada 20ms (1/50) enquanto, nos EUA, as televisões eram mais rápidas, podendo apresentar uma imagem a cada 16.7ms (1/60). Os monitores, seguindo a tecnologia das televisões, geralmente funcionam a 60Hz.

Embora consumam muita energia, os Monitores CRT possuem um excelente tempo de resposta aos comandos do computador (na casa dos microssegundos, μs), e excelente ângulo de visão, o que permite que pessoas em diferentes posições tenham uma experiência psicofísica da imagem muito semelhante. Por esta razão, diversos centros importantes de ciências cognitivas resistem às tecnologias recentes e insistem em apresentar estímulos visuais exclusivamente em monitores CRT.

5.3 Os monitores modernos são uma boa opção?

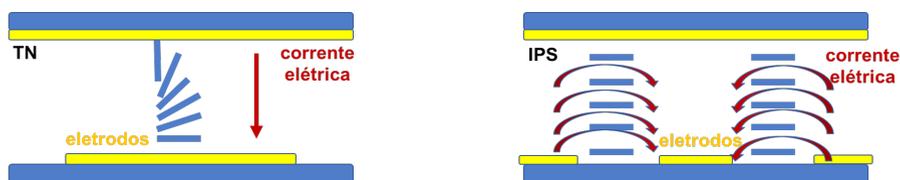
A tecnologia dos monitores mais modernos evoluiu a partir das telas monocromáticas utilizadas, por exemplo, em relógios e em alguns laptops antigos. Chamamos normalmente esta tecnologia de LCD (*Liquid Crystal Display*; Tela de Cristal Líquido) mas, o nome correto seria TFT (*Thin Film Transistor*, Transistor de Filme Fino). O LCD é apenas a forma como os primeiros monitores de tela fina funcionam.

O cristal líquido é uma substância transparente mas, ao receber a corrente elétrica, desmonta sua estrutura e se torna opaca, bloqueando passagem de luz. Nos monitores TFT-LCD, o cristal líquido é espalhado entre duas lâminas transparentes e polarizadas em sentidos opostos (HOOGBOOM *et al.*, 2007). Para formar a imagem, o transistor emite uma corrente elétrica capaz de alterar a configuração do LCD, fazendo as moléculas girarem até 90° na vertical. Por este motivo os monitores que utilizam esta tecnologia são chamados *Twisted Nematic* (LCD-TN), devido ao arranjo torcido das moléculas de cristal líquido que se posicionam de forma perpendicular à tela (Figura 6). O movimento das moléculas de cristal guia os raios de luz na formação da luz e das cores, de acordo com a imagem a ser exibida.

Algumas das vantagens do LCD-TN foi ter diminuído o tamanho dos monitores, ter um tempo de resposta ainda razoável (poucos milissegundos) e ser extremamente barato hoje em dia. Por outro lado, seu ângulo de visão é bastante restrito devido à angulação das moléculas de cristal líquido. Isso resulta numa baixa fidelidade de cores, de brilho e de contraste da imagem exibida. Estas características fazem com que monitores LCD-TN não sejam recomendados para estimulação visual visto que é difícil que dois participantes tenham a mesma experiência psicofísica da imagem. Nos LCD-TN, brilho, cores e contraste se alteram drasticamente bastando um sutil movimento para o lado.

Na busca por uma solução para este ponto fraco, foi elaborada a tecnologia LCD-IPS (*In-Plane Switching*) que, através de um novo método, conseguiu com que as moléculas de cristal líquido girassem no sentido horizontal ao invés do vertical das telas TN, se posicionando de forma paralela à tela. Esta mudança diminui a distorção da imagem e aumenta seu ângulo de visão. Monitores IPS possuem excelentes ângulos de visão e fidelidade de cores. Seu ponto fraco, porém, está em seu tempo de resposta que é bem mais lento que o dos monitores TN. Inicialmente este se tornou um dos grandes pontos fracos da tecnologia, criando o efeito *Ghosting*, quando percebemos resquícios das imagens anteriores nas imagens correntes devido a baixa taxa de atualização. A tecnologia IPS ainda não resolveu este problema mas, hoje, por falta de opções, ainda é o monitor LCD, de preço acessível, mais recomendado para experimentação (Figura 6).

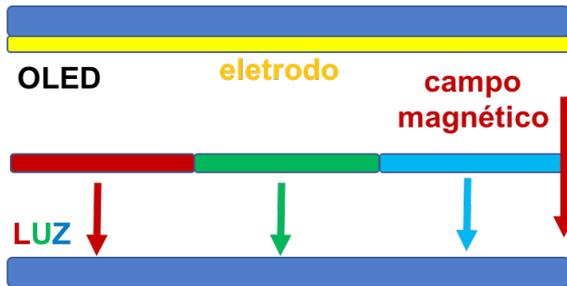
FIGURA 6 – Modelos de funcionamento de monitores LCD-TN e LCD-IPS



Nota: Adaptado dos manuais da Sharp.

Com a escassa oferta de monitores CRT no mercado, alguns trabalhos discutem a possibilidade de monitores OLED (*Organic Light-Emitting Diode*) e os displays de painel de plasma o substituírem (ITO *et al.*, 2013; RICHLAN *et al.*, 2013). No que se refere ao OLED, estes monitores são construídos com duas ou três camadas de materiais de carbono que emitem luz quando expostos a um campo eletromagnético (Figura 7). A primeira camada é responsável pela condução da energia elétrica enquanto a última é responsável pela emissão de luz. Esta luz é produzida por três lâminas, responsáveis pelas cores do sistema RGB (vermelho, verde e azul) e o brilho da luz é proporcional à força do campo magnético.

FIGURA 7 – Modelo de funcionamento de um monitor OLED



Nota: Baseado no modelo da Visionox.

Monitores OLED possuem gama e fidelidade de cores superiores, além de possuírem melhor brilho e um excelente campo de visão, girando em torno de 170-180 graus, o que evita as distorções comuns nos monitores LCD. Além disso, possuem um tempo de resposta também superior, o colocando à frente da concorrência em termos de usabilidade para a experimentação visual. Sua grande desvantagem é que se trata de uma tecnologia recente e, por isso, ainda possuem um alto custo no mercado. Cooper *et al.* (2013), em especial, já apontam os monitores OLED como ótimos substitutos aos CRT.

5.4 Questões sobre a tecnologia dos monitores, desenho e aplicação de testes psicolinguísticos

Atualmente, a grande maioria dos monitores funcionam a 60Hz, embora seja possível encontrar monitores de tecnologia mais recente com *refreshrates* de até 200Hz (5ms/quadro), especialmente para jogos. Ainda assim, a experiência de vídeo depende também da capacidade de o hardware processar cada imagem dentro desta taxa de atualização, através de um processador razoável, uma boa placa de vídeo, preferencialmente dedicada, e também uma quantidade razoável de memória RAM livre. Também é importante que, na hora da aplicação dos testes, sejam desabilitadas tarefas de fundo como antivírus, atualizações de software e notificações, a fim de evitar o consumo desnecessário de memória e de recursos do processador que podem deixar as tarefas mais lentas, alterando tanto a percepção quanto a medição dos dados.

Além disso, como indicado anteriormente, se precisamos de exatidão na casa dos poucos milissegundos, é de extrema importância termos noção sobre a competência e o desempenho da máquina. Para dar um exemplo do porquê, me basearei em um protocolo experimental de priming encoberto, no qual o experimentador apresenta a palavra *prime* por poucas dezenas de milissegundos a fim de que a palavra seja lida, mas dificilmente percebida pelo participante.

Considere um experimento de *priming* encoberto de Garcia (2013), em que a experimentadora quis apresentar a palavra *prime* por 38ms encoberta por uma máscara (uma sequência de *) que é apresentada por 50ms antes e após a palavra *prime*. Agora considere que este teste seja aplicado em um monitor de 60Hz. Ao utilizar um software com interface gráfica como o Psyscope (utilizado pela autora) ou o Open Sesame, um experimentador inexperiente irá indicar o tempo de apresentação desejado para o estímulo no campo correspondente: 38ms. Porém, isso quer dizer apenas que, aos 38ms, o computador enviará o comando para o monitor atualizar a imagem. Na prática, o estímulo só será de fato trocado no tempo de atualização da tela seguinte, ou seja, em um múltiplo de 16.7ms (1s / 60Hz). Neste cenário, podemos perceber que a última atualização do monitor teria sido por volta dos 33.4ms, o tempo de 2 quadros. Assim, a próxima atualização será aos $33.4\text{ms} + 16.7 = 50.1\text{ms}$, o que indica que tanto a palavra *prime* (38ms) quanto a sua máscara (50ms) seriam, na verdade, apresentados com a mesma duração, de 50.1ms.

A princípio, isso não é necessariamente um problema para grande parte dos experimentos, quando trabalhamos com períodos relativamente longos como 300 ou 400ms. Ainda assim é importante estar atento a esta questão uma vez que, no caso do *priming* encoberto citado e de outros testes cujos estímulos sejam apresentados por poucos quadros, o tempo de uma atualização da tela pode ser a diferença entre o participante ter consciência do estímulo ou não. Além disso, o experimentador irá descrever o teste indicando que a palavra *prime* foi apresentada por um determinado período quando, por limitações de hardware, ela teria sido apresentada por um tempo consideravelmente maior. Caso as configurações de hardware e software não estejam explicitamente citadas na seção de métodos, não é possível ter certeza de que a descrição do teste corresponde, de fato, a sua realidade.

O problema pode ser ainda mais grave. Enquanto muitos guias de uso especificam que seus monitores funcionam a 60Hz, não é raro

encontrar em suas especificações técnicas detalhadas a informação de que eles funcionam em um range entre 60-75Hz. E estas especificações não deixam claro em que situações o monitor funciona em determinada faixa, se a atualização possui um timing variável a depender do tipo de imagem ou se ele permite configurações que nos possibilite controlar o *refreshrate* dentro desta faixa. Diversos trabalhos sobre métodos em estimulação visual se debruçam sobre este tema. A maioria concorda que, (i) ao se trabalhar com monitores LCD, é necessário realizar testes de precisão e que (ii) confiar apenas na taxa de atualização e na contagem de quadros não é um método confiável (PLANT; TURNER, 2009; ELZE, 2010a, b; BAUER, 2015).

Quem realmente programa está atento a estas questões pois saber a taxa de atualização do monitor é essencial para que seu código funcione. Isso pode ser observado no código matlab/psychtoolbox desenvolvido por Sampaio & van Wassenhove (2013), utilizado e publicado por Sampaio (2015) e ilustrado na figura 8. Neste código, existe um cálculo em cima do número de frames apresentado pelo hardware (variável “*dur_f*”) para, enfim, adaptar o tempo indicado (variável “*dur*”) e reportar o número de frames apresentados. Este tipo de cálculo é comumente chamado de sincronizador adaptativo (*adaptive synchronizer*). Para usuários de softwares GUI, porém, estes detalhes podem passar uma vida sem serem percebidos.

FIGURA 8 – Captura de tela com parte do código escrito em Matlab por Sampaio & van Wassenhove (2013), utilizando funções do PTB 3

```
%%
fRate = FrameRate(0);
%
dur = 200; % ms
dur_f = fix(dur*fRate/1000); % frames
ITI_ms = [500,1000];

%% TRIALS
vlist = {'abafar','abaixar','abalar','abanar','abandonar','abastecer','abater','abrir','acalmar','acariciar',
tempoa = {'1: INSTANTE','2: SEGUNDOS','3: MINUTOS','4: HORAS','5: DIAS+'};
emotion = {'1: BOM','2: NEUTRO','3: RUIM',' '};

%%
nV = length(vlist); % total # of verba
nB = 1; % total # of blocks
ITI = fix(ITI_ms*fRate/1000); % range of ITI and ITI

for k = 1:nB
    tmpmat(:,2) = Shuffle(Shuffle(1:nV)); % COL 2 = word index in vlist
    tmpmat(:,1) = Shuffle(Shuffle(1:nV)); % COL 1 = trial number
    tmpmat = sortrows(tmpmat);
    expMat(k) = tmpmat;
    clear tmpmat
end
save(initials, '_itime_params_A')
end
```

O PsychoPy também possui um sincronizador adaptativo utilizando uma função para testar a taxa de atualização do monitor e calcular a duração de cada frame, como ilustrado na figura 9. Ainda assim, é importante estar atento pois, em alguns casos, ele não consegue recuperar a taxa de atualização e usará 60Hz como padrão. O E-Prime 2, possui

uma ferramenta de diagnóstico que também realiza esta sincronização e nos dá informações sobre a capacidade do hardware (SCHNEIDER *et al.*, 2002). Os desenvolvedores do Open Sesame, em seus tutoriais em vídeo,¹⁷ chamam atenção para este ponto ao recomendar a utilização de múltiplos da taxa *fps* decrescidos de 5ms, para evitar possíveis atrasos no processamento. Já o Paradigm e outros softwares conseguem reportar o número de quadros apresentados e, por vezes, a sua duração. Porém não fica claro se eles possuem algum tipo de sincronizador.

FIGURA 9 – Sincronizador adaptativo nativo do PsychoPy

```
# store frame rate of monitor if we can measure it
expInfo['frameRate'] = win.getActualFrameRate()
if expInfo['frameRate'] != None:
    ... frameDur = 1.0 / round(expInfo['frameRate'])
else:
    ... frameDur = 1.0 / 60.0 # could not measure, so guess
```

O Pyscope também não deixa claro se possui este sincronizador, mas Cohen & Provost (1994, p. 446) indicam a existência de outro método de controle, o *retrace synching*. Normalmente o computador envia a nova imagem para o monitor de acordo com a taxa de atualização. O monitor, então, aguarda até que o frame seja finalizado para atualizar a imagem, como vimos anteriormente. Com o *retrace synching*, o PsyScope aguarda o sinal de retorno do monitor (*retrace signal*) para enviar a imagem, o que garante que o tempo indicado nos resultados é o tempo exato do *onset* do estímulo. A duração exata dos frames e, por consequência, do estímulo, pode ser calculada a partir dos dados do *log* e do *retrace synching*.

Todas estas questões nos mostram que, para desenvolver testes psicofísicos e psicolinguísticos, não basta ter domínio sobre um determinado software. É necessário também ter alguma noção sobre que passos o hardware deverá seguir e o quanto bem ele é capaz de executar estes passos, de forma que possamos pensar em como contornar eventuais problemas ou desvios de precisão e de exatidão os dados. Softwares com interface gráfica, são bastante úteis por simplificar a tarefa de elaborar um experimento. Por outro lado, eles nos permitem rodar experimentos sem a

¹⁷ www.youtube.com/ceebassmusic

necessidade de entender o que, de fato, está sendo feito das variáveis que definimos para o teste. Isso pode nos fazer acreditar que uma determinada variável visual ou temporal está devidamente controlada quando, na verdade, não está.

Após esta discussão sobre testes rodando numa máquina controlada, me pergunto sobre a precisão e exatidão dos dados retornados em ferramentas web que não possuem máquinas, monitores e nem ambientes controlados.

6. Considerações Finais

Ao final deste trabalho, acredito ter atingido dois objetivos principais. O primeiro deles é a discussão e apresentação de diversos tipos e opções de software que podem ser utilizados para experimentação em ciências cognitivas em geral, o que inclui a psicolinguística. Existe uma enorme gama de softwares em diferentes plataformas e com diferentes níveis de curva de aprendizagem que poderiam ser muito mais difundidos no Brasil, aumentando o contato dos alunos de Linguística com a experimentação. O segundo objetivo é a discussão sobre eventuais problemas de método que podem ser facilmente contornáveis caso tenhamos conhecimento do que acontece na máquina quando estamos rodando nosso teste.

6.1 Mas qual software eu devo utilizar?

Uma das principais perguntas que poderá ser feita após esta discussão é: “qual software devo utilizar”? Acredito que minha contribuição neste artigo foi a de apresentar diversas opções e suas principais características, de forma que você tenha alguma base antes de escolher um deles. De uma forma mais prática, apesar de antigamente o DMDX, o Presentation e o Psyscope serem alguns dos mais utilizados, hoje percebo que os softwares mais populares são o C e suas toolboxes, o PyGame (Python) e o Psychtoolbox 3 (Matlab) entre aqueles que programam. Entre os que não programam, o E-Prime e o PsychoPy me parecem ser os mais populares nos laboratórios americanos, europeus. No Brasil, o E-Prime se tornou bastante popular nos últimos anos entre os não-programadores, seguido pelo Paradigm, devido ao seu preço mais acessível. Dentre as opções em software livre, vejo raros artigos utilizando o PsychoPy e o DMDX.

Para quem inicia sua vida acadêmica, acredito ser bastante razoável recomendar o PsychoPy. Esta recomendação se deve a 5 fatores: (i) se trata de um software extremamente simples com uma interface gráfica limpa, que contem apenas o necessário; (ii) te dá a possibilidade de continuar usando o mesmo software após uma iniciação à programação em Python, (iii), possui um sincronizador adaptativo, te dando maior confiança quanto aos dados obtidos; (iv) é bastante popular e você poderá trocar experiências e códigos com diversos pesquisadores do mundo que o utilizam, além de (v) se tratar de software livre. Vale ressaltar que esta indicação não passa de uma opinião pessoal de um software que considero ser extremamente prático e confiável para grande a maioria dos casos.

Particularmente, tenho uma ótima experiência ao utilizar o PsychoPy em aulas de psicolinguística para a graduação. Esta experiência faz com que os alunos percam o medo de elaborar experimentos por não saberem programar, consigam aplicar e analisar seus próprios testes em poucas aulas e, por consequência, tenham uma experiência mais real sobre o que é a experimentação psicolinguística, aumentando o interesse pela área. Além disso, os laboratórios de Psicolinguística no Brasil costumam pagar mais de mil dólares em cada licença que podem ser facilmente substituíveis por soluções em software livre, bastando um pouco mais de informação. Apesar de livres, todos estes softwares possuem grupos de discussão que funcionam como um suporte coletivo entre os usuários e os desenvolvedores responsáveis.

Além do PsychoPy, as opções em JavaScript parecem ser excelentes opções para coleta de dados via web. Embora eu ainda não me sinta a vontade para realizar coleta de tempos de resposta nestas plataformas, as comparações de De Leeuw & Moritz (2015) me pareceram consistentes. Me pergunto apenas se a precisão cronométrica se mantém independente da diferença de processamento das máquinas utilizadas e dos seus dispositivos. Por esta razão, apesar de recomendar, sugiro um certo cuidado com estas plataformas caso você trabalhe com diferenças sutis na estimulação física, como diferenças em imagens, em intensidade de luz e de cores ou com tempos de apresentação como nos testes de priming encoberto.

6.2 Atenção às configurações de software e de hardware

Meu segundo objetivo foi mostrar que é necessário ter uma compreensão mínima da interface software-hardware para que possamos ter certeza de que controlamos corretamente a estimulação psicolinguística. Apenas dizer ao computador o que queremos não significa que ele é capaz de realizar. Sem um conhecimento das capacidades do hardware ou sem a utilização de medidores externos acurados, é impossível perceber que a máquina não está controlando os tempos da forma que indicamos.

Ainda neste escopo, estas questões mostram a importância de descrevermos detalhadamente o software, código e hardware utilizados no desenho e aplicação dos testes. Muitos softwares podem não ter sido testados em uma determinada versão de um sistema operacional, especialmente os recém lançados. Por esta razão, não podemos atualizar os sistemas operacionais das plataformas de aplicação antes de termos certeza de sua total compatibilidade com os softwares, o que é diferente de o software simplesmente funcionar.

Alguns softwares podem apresentar problemas com determinadas peças de hardware como, por exemplo, uma placa de vídeo, mas raramente estamos atentos aos avisos dos desenvolvedores sobre estas questões. Além disso, muitas vezes indicamos em nossos testes que a apresentação dos estímulos foi realizada em um tempo que é notadamente impossível de ser apresentado em um monitor comum. Isso não é errado visto que não temos a obrigação de conhecer todos os detalhes, configurações e incompatibilidades de hardware existentes. E por isso é importante sinalizar estes detalhes em nossos métodos de forma que possíveis problemas possam ser facilmente identificados por aqueles que possuem um maior conhecimento no assunto. Esses cuidados evitam alguns dos fatores que levam ao problema da replicação dos resultados, tema que vem sendo bastante debatido como, por exemplo, no levantamento da Open Science Collaboration (2015) na Science que levou, posteriormente, à publicação de “*A manifesto for reproducible science*” (MUNAFÒ *et al.*, 2017), na Nature.

Reforço que a divulgação detalhada e cuidadosa das principais informações sobre o hardware, o software, a elaboração, os métodos de elaboração, de aplicação e de análise de nossos experimentos, assim como a divulgação dos estímulos e dos códigos fonte, são fatores fundamentais para a viabilidade do método experimental, que tem sua eficácia e validade fundadas exatamente na reprodução sistemática destes

métodos e de seus resultados por diferentes pesquisadores em diversas partes do mundo.

Acredito que este artigo seja de alguma forma inspirador para que sejamos mais atentos e tenhamos um maior cuidado descritivo no momento de reportar nossos testes.

Agradecimentos

Agradeço a Virginie van Wassenhove, Douglas Bemis, Jansen Oliveira, Daniela Cid de Garcia e Julia Cataldo Lopes pelas discussões sobre linguagem de programação, experimentação e métodos. Renata Passetti que, há poucos dias do envio deste artigo, me inspirou a revisar as opções de experimentação via web, tornando este trabalho um pouco mais abrangente. Agradeço a Leticia Kolberg e aos alunos do curso Tópicos em Psicolinguística em 2016/2 que, nas aulas sobre métodos, me ajudaram a levantar alguns dos exemplos utilizados ao longo do texto. Agradeço aos revisores deste artigo que, com suas sugestões de reelaboração, tornaram este trabalho, na medida do possível, um pouco mais acessível. O presente artigo foi realizado sob a vigência do projeto FAEPEX 519.292 e no âmbito do auxílio FAPESP 2016/13.920-9.

Referências

BAUER, B. A Timely Reminder About Stimulus Display Times and Other Presentation Parameters on CRTs and Newer Technologies. *Canadian Journal of Experimental Psychology*, Société Canadienne de Psychologie, v. 69, n. 3, p. 264-273, 2015. <https://doi.org/10.1037/cep0000043>.

BEZANSON, J.; EDELMAN, A.; KARPINSKI, S.; SHAH, V.B. Julia: a fresh approach to numerical computing. *ArXiv*, 2014. Disponível em: <arxiv.org/abs/1411.1607>. Acesso em: 28 nov. 2016.

BRAINARD, D. H. The psychophysics toolbox. *Spatial Vision*, Brill Online, n. 10, p. 433-436, 1997.

COHEN, J.; MACWHINNEY, B.; FLATT, M.; PROVOST, J. PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments & Computers*, Springer Link, v. 25, n. 2, p. 257-271, 1993.

COHEN, J.; PROVOST, J. *PsyScope: User Manual 1.0*, Carnegie Mellon University, 1994. Disponível em: <psy.cns.sissa.it/psy_cmu_edu/PsyMan.pdf>. Acesso em: 28 nov. 2016.

COOPER, E. A.; JIANG, H.; VILDAVSKI, V.; FARRELL, J. E.; NORCIA, A. M. Assessment of OLED displays for vision research. *Journal of Vision*, Association for Research in Vision and Ophthalmology, v.13, n. 16, p. 1-12, 2013.

DE LEEUW, J. R. jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, Springer, v. 47, n.1, 1-12, 2014. <https://doi.org/10.3758/s13428-014-0458-y>.

DE LEEUW, J. R.; MOTZ, B. A. Psychophysics in a Web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behavior Research Methods*, Springer, v. 48, n.1, p.1-12, 2015. <https://doi.org/10.3758/s13428-015-0567-2>.

ELZE, T. Misspecifications of Stimulus Presentation Durations in Experimental Psychology: A Systematic Review of the Psychophysics Literature. *PLoS ONE*, São Francisco, Califórnia, v. 5, n. 9, 2010a.

ELZE, T. Achieving precise display timing in visual neuroscience experiments. *Journal of Neuroscience Methods*, Elsevier, n. 191, p. 171-179, 2010b.

FORSTER, K. I.; FORSTER, J. C. DMDX: A Windows display program with millisecond accuracy. *Behavioral Research Methods*, Springer, n. 35, p. 116-124, 2003.

GARAIZAR, P.; VADILLO, M.A.; LÓPEZ-DE-IPÍÑA, D.; MATUTE, H. Measuring Software Timing Errors in the Presentation of Visual Stimuli in Cognitive Neuroscience Experiments. *PLoS ONE*, São Francisco, Califórnia, v. 9, n. 1, 2014.

GARCIA, D.C. *Elementos estruturais no acesso lexical: o reconhecimento de palavras multimorfêmicas no português brasileiro*. 2009. 108 f. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2009.

HOOGBOOM, J.; ELEMANS, J. A. W.; ROWAN, A. E.; RASING, T. H. M.; NOLTE, R. J. M. The development of self-assembled liquid crystal display alignment layers. *Philosophical Transactions of The Royal Society A*, The Royal Society Publishing, n. 365, p. 1553-1576, 2007.

ITO, H.; OGAWA, M.; SUNAGA, S. Evaluation of an organic light-emitting diode display for precise visual stimulation. *Journal of Vision*, Association for Research in Vision and Ophthalmology, v. 13, n. 7, p. 1-21, 2013.

KENKEL, F. Untersuchungen über den Zusammenhang zwischen Erscheinungs-grobe und Erscheinungsbewegung bei einigen sogenannten optischen Tauschungen. 2. *Zeitschrift für Psychologie*, Göttingen, v. 67, p. 358-449, 1913.

KLEINER, M.; BRAINARD, D.; PELLI, D. What's new in Psychtoolbox-3? *Perception*, v. 36, n. 14, p. 1-26, 2007.

KRAUSE, F.; LINDERMANN, O. Expyriment: A Python library for cognitive and neuroscientific experiments. *Behavior Research Methods*, Springer, v. 46, n. 2, p. 416-428, 2014. <https://doi.org/10.3758/s13428-013-0390-6>.

LANGE, K; KÜHN, S.; FILEVICH, E. Just Another Tool for Online Studies (JATOS): An easy solution for setup and management of web servers supporting online studies. *Plos One*, São Francisco, Califórnia, v. 7, n. 10, 2015.

MATHÔT, S.; SCHREIJ, D.; THEEUWES, J. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavioral Research Methods*, Springer, v. 44, n. 2, p. 314-324, 2012. <https://doi.org/10.3758/s13428-011-0168-7>.

MEDINA J. M.; WONG, W.; DÍAZ, J. A.; COLONIUS, H. Advances in Modern Mental Chronometry. *Frontiers in Human Neuroscience*, Frontiers, v. 9, n. 256, p. 5-7, 2015. <https://doi.org/10.3389/fnhum.2015.00256>.

MOORE, G. E. Cramming more components onto integrated circuits. *Electronics*, v. 38, n. 8, p. 114-117, 1965.

MUELLER, S.T.; PIPER, B.J. The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods*, Elsevier, n. 222, p. 250-259, 2014.

MUNAFÒ, M. R.; NOSEK, B. A.; BISHOP, D. V. M.; BUTTON, K. S.; CHAMBERS, C. D.; DU SERT, N. P.; SIMONSOHN, U.; WAGENMAKERS, E. J.; WARE, J. J.; IOANNIDIS, J. P. A. A manifesto for reproducible science. *Nature Human Behaviour*, Springer Nature, n.1, 2017.

OPEN Science Collaboration. Estimating the reproducibility of psychological science. *Science*, American Association for the Advancement of Science, v. 349, n. 6251, 2015.

PEIRCE, J. W. PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, Elsevier, v. 162, n. 1-2, p. 8-13, 2007. <https://doi.org/10.1016/j.jneumeth.2006.11.017>

PEIRCE, J.W. Generating stimuli for neuroscience using PsychoPy, *Frontiers in Neuroinformatics*, Frontiers, v. 2, n. 10, 2009.

PLANT, R. R.; HAMMOND, N.; TURNER, G. Self-validating presentation and response timing in cognitive paradigms: How and why? *Behavior Research Methods, Instruments, & Computers*, Springer Link, n. 36, p. 291-303, 2004.

PLANT, R. R.; TURNER, G. Millisecond precision psychological research in a world of commodity computers: New hardware, new problems? *Behavior Research Methods*, Springer, v. 41, n. 3, p. 598-614, 2009. <https://doi.org/10.3758/BRM.41.3.598>.

READ, P.; MEYER, M. P. *Restoration of motion picture film*. Oxford: Butterworth-Heinemann, 2000. (Series in Conservation and Museology)

REIMERS, S.; STEWARD, N. Presentation and response timing accuracy in Adobe Flash and HTML5/JavaScript Web Experiments, *Behavior Research Methods*, Springer, v. 47, n. 1, p. 309-327, 2014.

RICHLAN, F.; GAGL, B.; SCHUSTER, S.; HAWELKA, S.; HUMENBERGER, J.; HUTZLER, F. A new high-speed visual stimulation method for gaze-contingent eye movement and brain activity studies. *Frontiers in Systems Neuroscience*, Frontiers, v. 7, n. 24, 2013.

SAMPAIO, T. O. M.; VAN WASSENHOVE, V. Self-paced Reading tests for GNU Octave/Matlab [software computacional], 2013. Disponível em: <http://www.thiagomotta.net/uploads/7/0/5/2/7052840/spr_tests_-_octave-matlab_13.zip>. Acesso em: 30 mar. 2017.

SAMPAIO, T. O. M. *Coerção aspectual: uma abordagem linguística da percepção do tempo*. 2015. 398f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2015.

SCHNEIDER, W.; ESCHMAN, A.; ZUCCOLOTTO, A. E-Prime user's guide. Pittsburgh, PA: Psychology Software Tools, 2002. Disponível em: <step.psy.cmu.edu/materials/manuals/users.pdf>. Acesso em: 28 nov. 2016.

SHINNERS, P. *PyGame - Python Game Development* [computer software], 2011.

STOET, G. PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, Springer, v. 42, n. 4, p. 1096-1104, 2010. <https://doi.org/10.3758/BRM.42.4.1096>.

STRAW, A. D. Vision Egg: An Open-Source Library for Realtime Visual Stimulus Generation. *Frontiers in Neuroinformatics*, Frontiers, v. 2, n. 4, 2008. <https://doi.org/10.3389/neuro.11.004.2008>.

TURPIN, A.; LAWSON, D.J.; MCKENDRICK, A.M. PsyPad: a platform for visual psychophysics on the iPad. *Journal of Vision – Methods*, The Association for Research in Vision and Ophthalmology, v. 14, n. 16, 2014.

WATSON, A. B. *Handbook of Perception and Human Performance*. New York: Wiley, 1986.

WERTHEIMER, M. Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, Göttingen, v. 61, n. 1, 161-265, 1912.

ANEXO: Endereço das ferramentas computacionais citadas no texto**a) Multiplataforma Livre:**

C: www.open-std.org/jtc1/sc22/wg14

ExPyrimint: www.expyriment.org

JATOS: www.jatos.org

Java: www.java.com

JsPsych: <http://www.jspsych.org>

Julia: julialang.org

Open Sesame: osdoc.cogsci.nl

Octave: www.gnu.org/software/octave

PEBL: pebl.sourceforge.net

PsyToolKit: www.psytoolkit.org

PsychJava: psychjava.com*

PsychoPy: psychopy.org

Psychtoolbox 3 (p/ Matlab e Octave): psychtoolbox.org

PsyPad: www.psympad.net.au

PyGame: pygame.org

Python: www.python.org

R-Project: www.r-project.org

Scilab: www.scilab.org

VisionEgg: visionegg.org

b) Multiplataforma Proprietário:

Matlab: www.mathworks.com

SuperLab: www.cedrus.com/superlab

c) MacOs X, Livre:

PsyScope: psy.ck.sissa.it

d) Windows, Proprietários:

E-Prime: www.pstnet.com/eprime.cfm

Paradigm: paradigmexperiments.com

Presentation: www.neurobs.com

e) RaspberryPi: www.raspberrypi.org

* O site do PsychJava está fora do ar há algum tempo. Não consegui informações sobre os motivos da queda do site. Como ele já estava incorporado em outros softwares, acredito que o projeto esteja parado.

Maxakalí Nasality and Field Recording with Earbud Microphony

Nasalidade no Maxakalí e gravação em campo com microfonia de fone de ouvido

Andrew Ira Nevins

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
a.nevins@ucl.ac.uk

Mário André Coelho da Silva

Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais / Brasil
mario.andrecs@gmail.com

Abstract: The challenges in verifying vowel nasalization acoustically are well-known, and for this reason many researchers have opted for the use of nasal airflow masks. Such equipment, however, is not always affordable for recordings or projects conducted in the field. In the present article, we apply the recently developed technology of earbud recordings in Stewart & Kohlberger (2017) to the study of nasality in Maxakalí as a new method in experimental fieldwork phonology. This language has several phenomena related to nasality such as nasal spreading, prenasalization at the beginning of words, and glottal transparency in morphologically-alternating ‘long forms’. One goal in using methodology with affordable equipment is to make it easier to generate a visual representation of the timing of the nasal profile of a word, thereby yielding possibilities for experimental linguistics to show increased integration with language description and pedagogy.

Keywords: nasal harmony; prenasalized stops; loanword phonology; Maxakalí; field recording.

Resumo: As dificuldades em se verificar acusticamente a nasalização de vogais são bem conhecidas e por isso diversos pesquisadores optam por usar máscaras de fluxo nasal. Porém, esse tipo de equipamento nem sempre é prático para gravações em campo. No presente estudo, foi aplicada a tecnologia recentemente desenvolvida de gravação com fones de ouvido apresentada em Stewart e Kohlberger (2017), com o intuito de estudar a nasalidade no Maxakalí e esse processo de gravação como um novo método de fonologia experimental em campo. A língua Maxakalí conta com diversos fenômenos ligados à nasalidade como, por exemplo, espraçamento de nasalidade, pré-nasalização em início de palavra, entre outros. Um objetivo ao usar uma metodologia com equipamentos de baixo custo é facilitar a geração de uma representação visual do perfil nasal de uma palavra, produzindo possibilidades para a linguística experimental mostrar uma integração maior com a pedagogia de língua.

Palavras-chave: fonologia de empréstimos; gravação em campo; harmonia nasal; Maxakalí; oclusivas pré-nasalizadas.

Recebido em: 10 de dezembro de 2016.

Aprovado em: 3 de abril de 2017.

1 Background on Maxakalí and difficulties of nasality in Acoustic Analysis

1.1 An overview of Maxakalí phonology

Maxakalí is an indigenous Brazilian language spoken in Minas Gerais state by about 2,000 people. It forms part of the Macro-Jê stock and it is the only extant language within its branch. Maxakalí is well known in the phonological literature for a number of properties such as the lack of any phonological liquids and the prevocalization of all its eight consonants in coda position (GUDSCHINSKY; POPOVICH; POPOVICH, 1970; WETZELS, 1993; WETZELS; SLUYTERS, 1995; ARAÚJO, 2000; SILVA 2015). The table below shows the language's consonantal inventory, as well as the prevowels resulting from coda lenition (or prevocalization):

	labial	dental	palatal	velar	glottal
voiceless stop	p [p̥]	t [t̥]	c [ç]	k [k̟]	ʔ
voiced / nasal stop	b / m [b̥]	d / n [d̥]	ʝ / ɲ [j̥]	g / ŋ [g̥]	
fricative					h

An important allophonic process in Maxakalí is the one involving voiced and nasal stops: nasal stops always precede nasal vowels and voiced stops most of the times precede oral vowels. The two instances where voiced stops precede nasal vowels are in some verbs with the [-nãŋ] suffix (further discussed in Section 3.5) and in loanwords. It is also important to know that the coda position always agrees in nasality with the syllable nucleus, so the consonant in this position is underspecified for nasality (SILVA; NEVINS, 2015). As can be seen from the Table above, glottals do not undergo lenition, as they never occur in coda position (except for some loanwords). Therefore, both /ʔ/ and /h/ do not have a prevocalized allophone.

There is a long discussion in the literature about the underlying nature of voiced ~ nasal stops. Gudschinsky, Popovich & Popovich (1970) contend that Maxakalí has nasal stops and a contrast between oral and nasal vowels, such that nasals become oral before oral vowels (as represented in Popovich's orthography). Rodrigues (1981) claims that, except for a few instances, there are just oral stops and oral vowels in the language and that by the means of a rule, codas become nasalized and then spread nasality to the segments to their left. Araújo (2000) argues that there are nasal consonants and oral vowels only, and that nasal coda consonants are responsible for spreading nasality to other segments. Finally, Wetzels (2009) postulates that there are oral consonants only and both nasal and oral vowels. According to him, stressed nasal vowels are responsible for spreading nasality to other segments. In this paper, we follow Wetzels' hypothesis, as Gudschinsky, Popovich & Popovich's is typologically unnatural, Rodrigues' cannot explain why words without a coda may be nasal (then treating them as exceptions) and Araújo's is not adequate to explain why epenthetic vowels in CCVC words are nasal and not oral.

The syllabic structure of Maxakalí, according to Gudschinsky, Popovich & Popovich (1970), consists of a maximum of CVC, in which any consonant can be placed in onset position and, as said before, all consonants but the glottals can occupy coda position. Silva (2016) argues that the maximum syllable is, in fact CCVC – but this larger syllable type is restricted to cases in which the first onset consonant must be a non-coronal (labial or velar) and the second onset position must be filled in almost every case by a coronal. This structure corresponds to Proto Macro-Jê *C_rVC with a Proto-Maxakalí change of *_r > d > t.

Lexical stress falls on the last syllable of the word, but because of some interactions between the nuclear vowel and the coda consonant, it can be located in the penultimate syllable. This interaction has to do with the insertion of a glide, which can be predicted by the vowel and the prevowel resulting from consonant lenition. The table below shows which vowels interact with which coda consonant and some examples:

	dental	palatal
/i/	[j] /cit/ > ['tɛijə] 'thread'	-
/a/	[ɥ] /tat/ > ['tauɥə] 'get (uncount.)'	-
/i/	[ɥ] /-kít/ > ['kiɥə] 'louse'	[ɥ] /cic/ > ['tɛiɥi] 'leaf'
/o/	[w] /kot/ > ['kɔwə] 'tobacco'	[w] /poc/ > ['pɔwi] 'arrow'

The same phenomenon also can be observed with nasal vowels, as in /pĩd/ > ['pĩũĩ] *jump*, except for the interaction with /a/: /bãd/ > ['mãũ] and not *['mãũĩ] *shoot*. The final consonant, even though uncommon, can sometimes be heard in this context: ['kiɥũ] ~ ['kiɥũt]. For further discussion, see Wetzels (1993) and Wetzels & Sluyters (1995).

Some other phonological phenomena (besides nasal spreading, which, as mentioned, will be discussed in Section 3.5) found in Maxakalí regarding nasality are the prenasalization of consonants, left-edge nasality, glottal transparency and nasality in loanwords. The first of these, which will be presented in section 3.2, is found in words beginning with voiced stops, both in native and loanwords, although not in medial position.¹

¹ It seems that compound words where the second element begins with a voiced stop have some degree of variation (e.g.: /kik + bac/ > [kɛɥ 'baj] ~ [kɛɥ^m 'baj] 'turtle'). The pattern in compounds, however, awaits more extensive confirmation with recordings of the type we explore in the text (and a range of compound types).

- (1) /bai/ > [m̥ba'i] ~ [ba'i] 'to be good'
 (2) /kibik/ > [ki'biu], but *[ki'm̥biu] 'to be bad'

Somewhat related to this, there is left-edge nasality (section 3.1) in which vowels are nasalized from left to right, an apparent exception to the better understood right-edge nasality. There are just a few native words which fail to nasalize in this onsetless, word-initial context, and many loanwords have been adapted with nasal vowel even when in the source language there is no nasality, such as in (3):

- (3) Portuguese [oli'verə] > [õni'veɾɛ] 'Oliveira' (surname)

In section 3.3, we present evidence for glottal transparency. In Maxakalí, there are some nouns and verbs which alternate between short and long forms, according to syntactic and/or prosodic context. These words have a short form CVC and a long form CVHVC, in which the H corresponds to a glottal consonant and both Vs have the same phonetic value. We will argue that long forms are derived from the short ones, and that the glottal consonant cannot block nasal spreading, thus remaining transparent for nasality.

In section 3.4, we discuss nasality in loanwords from Portuguese. Some of the phenomena discussed above also apply to loanwords, but there are some exceptions, such as variation in nasal spreading. An example in which nasality varies is shown below:

- (4) [fla'mẽgo] > [pãñã'mãĩ] ~ [panã'mãĩ] ~ [padã'mãĩ] ~ [palã'mãĩ]
 'Flamengo' (soccer team)

In short, Maxakalí has prenasalization, coda nasalization, and nasal harmony within the syllable, between nucleus and rime, between rime and onset, and from syllable-to-syllable. Each of these processes show potential locality effects, based on factors such as loanword status, morphological restrictions, and intervening glottals. Development of theoretical models of the underlying forms and the nature of these processes therefore depends on a firm empirical basis for knowing the surface forms themselves.

1.2 Prior techniques for studying nasality

As the discussion above reveals, Maxakalí is a rich domain for the study of nasality, and many of its details present interesting challenges for theoretical models. Nonetheless, as is widely acknowledged, instrumental verification of nasality is not entirely straightforward. The literature contains varying suggestions for acoustic landmarks of nasality (such as CHEN, 1996; STYLER, 2015), but these are largely developed for vowel nasality, and as discussed above, many of the phenomena of interest in Maxakalí involve consonants as well. Measurements using nasal airflow masks (SHOSTED, 2006; DEMOLIN, 2011) are extremely informative, but such equipment is often expensive for researchers based in the ‘Global South’. To take an example that is all-too common, there are few if any Brazilian linguistics researchers or postgraduate students who have the budgetary support for equipment of this kind. Even costs aside, the analytic techniques involved in calibration and processing of aerodynamic data require technical training that may not even be available in all regions where these languages are being spoken and studied. And bringing a range of demographically diverse participants from a village setting into an urban laboratory hundreds of kilometers away is also not always practical, not only for financial but also for cultural reasons. On the other hand, earbuds are inexpensive and portable, and familiar to participants who have seen them with mobile phones, and we have found that they are welcome in virtually anyone’s home in the village, and are perceived as fairly discreet even in scenarios where the recording takes place among a range of extended family members of varying ages. Earbud recordings and the corresponding visualization can be informative in confirming or disconfirming impressionistic percepts of nasalization by fieldworkers and language teachers even with a limited background in acoustic phonetics. The main output of this technique is *nasalance*, the relative prominence of the amplitude of nasality versus orality, originally developed in Fletcher, Sooudi & Frost (1974) for clinical purposes, which can be graphed temporally with a corresponding transcription to show the details of timing of velar opening.

Nasalance was originally measured using an acoustic ‘baffle’, a plate separating the nostrils from the upper lip. However, the method we use follows Stewart & Kohlberger (2017)’s proposal of earbuds as a method of field recording, with no baffle in place. Prior acoustic studies using this method include Montagu & Amelot (2005) and Audibert &

Amelot (2011). Stewart & Kohlberger bring to the fore the usefulness of this technique in studying phenomena such as nasal harmony, which, given their geographic distribution, tend to cluster in regions within the Global South, where nearby training resources call for a more intuitive method for instantly displaying the action of the velum in controlling nasality. This is an invaluable exercise when working with native speakers to discuss oral/nasal contrasts as they relate to, say, orthographic marking (i.e. should prenasalization be marked? Well, we now have very easy methods to see how consistently it is produced – and whether the variation found is phonologically-predictable, lexically specific, or negligible, and thus making corresponding orthographic decisions). Given that all of the scripts for their paper are now openly available on the *Language Documentation and Conservation* journal site accompanying Stewart & Kohlberger (2017)'s publication, and that all of our recordings reported in this study are available with their corresponding segmented textgrids, any reader of this paper can analyze or visualize the entirety of the data we have collected. We strictly follow the method outlined in Stewart & Kohlberger (2017) for using earbuds as microphones to record a nasal track, alongside the simultaneous recording of an oral track and subsequent software post-processing as a means of representing nasal airflow, and focused our efforts on an experimental and demographic design and method that would allow us to collect a body of data and a set of illustrative analyses for well-known unresolved descriptive questions, the resolution of which leads itself to more well-informed theoretical and orthographic decisions.

2 The present study

2.1 Equipment and recording procedure

For data collection we used a Zoom h4n recorder, as it has two microphone inputs. The Zoom h4n recorder has the advantage of being able to record two tracks simultaneously, although two recorders (one for recording oral track and the other for nasal one) would be equally suitable for data collection. The most important thing is that these tracks be recorded separately.

In one of the inputs, we connected a wired CSR Pro 2.1 Microphone for the oral track recording. In the other input we plugged

Samsung earbuds with silicon tips (the price of a single pair, as of the time of writing, is around 50 BRL). A P2 x P10 adaptor was needed as the earbuds have a P2 connection, while the recorder has a P10/XLR input. The earbuds should preferably be of low impedance (~27 ohms):

Because earphones work on the same basic principle as microphones (i.e. a diaphragm vibrates when agitated by a medium and a copper coil around a magnet creates an electromagnetic field which interprets the signal), earbuds function as small microphones when attached to the input jack of a recording device. (STEWART; KOHLBERGER, 2017, p. 52).

With both microphone (for the oral track) and earbuds (for the nasal track) connected to the recorder, the participant was asked to hold the earbuds below each nostril, so nasal flow could be recorded. The oral track was recorded with the microphone, which was held by the researchers. At first, we tried a wireless mic on the oral track, so no one would have to hold the microphone. However, its wireless transmission generated low-level background noise that was picked up by the nasal earbuds, so we opted for a wired microphone. The downside of this wired microphone setting was that the recording procedure thereby needed one person to hold the microphone for the participant (whose own hands were busy with the two earbuds), and this time we did not take a tripod to the field.

Participants were asked for the translations of the words collected for this study, while holding the earbuds below their nostrils. The wordlist was randomized for each participant, in order to avoid any bias. The wordlist and the translation of its items can be found in Appendix A. As the answers in terms of translation equivalents were open, sometimes participants did not give the expected ones. For example, when asked for ‘doggy’ (Pt. *cachorrinho*), the expected Maxakali elicitation was for *kokexnãg* (dog-DIM), but a few participants said things like *kokex kutok* (dog-offspring) or *kokex kutĩynãg* (*little dog*). Thus not every participant provided the same set of tokens as every other.

2.2. Post-processing

Recordings were amplified at 12dB with Zoom’s built in amplifier on the nasal channels. We first selected the words of interest in Praat

in both tracks in parallel using the Group checkbox, and then made a textgrid based on the oral track, segmenting each recorded token into a phonetic form with the SAMPA transcription. We then used Praat to generate Intensity and Sound textfiles for the words of interest. All of these sound files can be found in the accompanying materials to this article, in the format axõk-oral-PM and axõk-nasal-PM, where these are the recording from the oral and nasal tracks, respectively, alongside a corresponding textgrid of the form ãxok-PM.textgrid (where PM, in this case, is a code for this particular Maxakali speaker). By superimposing the two intensity curves and using Stewart & Kohlberger's (2017) script (available at <http://nflrc.hawaii.edu/ldc/> as of April 2017) to normalize the two tracks using intensity in decibels, one can generate the visualized superimposed curves shown in the discussion below.

2.3 Demographics

We recorded 5 female speakers (DM, EM, MM, MSM, and SM) and 6 male speakers (GM, IM, IZM, PM, RM & TM), all of them living in Aldeia Verde (one of the four Maxakali villages), located near the town of Ladainha, Minas Gerais state. There are about 350 inhabitants in this village. People in virtually all households but one speak only Maxakali to each other. Most people in this village have at least some understanding of Portuguese, in a continuum of competence: some people are fluent in both Maxakali and Portuguese on the one hand, but on the other hand some inhabitants (mainly women) can understand but do not produce full sentences in Portuguese.

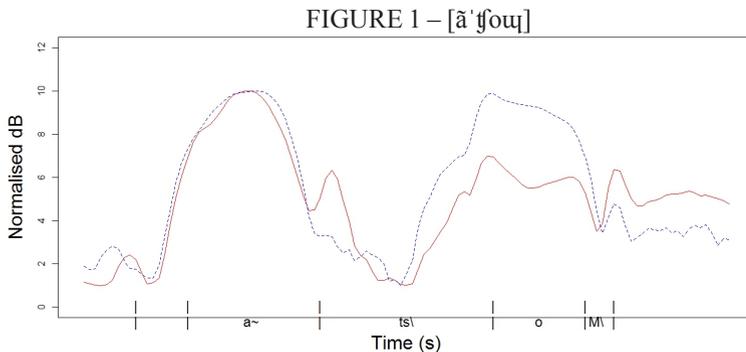
In terms of age, there were 6 younger speakers (DM, EM, GM, IM, and RM), with apparent age spanning from 15 to 30 years old, 4 adult speakers (MSM, SM, IZM, and PM) from approximately 30 to 50 years old, and 1 elder speaker (TM) more than 80 years old.

3 Empirical issues

In the following subsections, we highlight five empirical issues that have been the matter of some debate within the description of Maxakali phonology, and which receive a more concrete empirical picture given the ability to document nasality and clearly represent it with normalised intensity curves aligned with the segmental transcription of the accompanying token.

3.1 Left-edge nasality

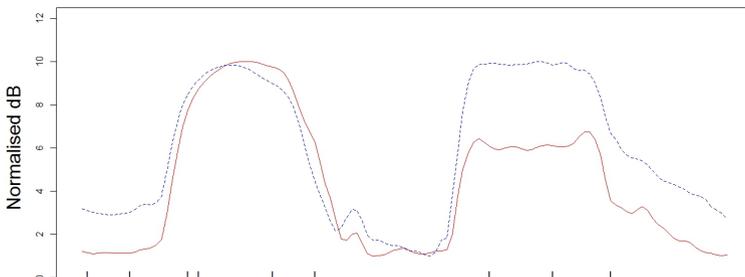
A phenomenon noted in Silva & Nevins (2015), independently noted in the transcriptions in Popovich & Popovich (2005), is the spontaneous nasalization of vowel-initial loanwords that have neither nasal vowels nor nasal consonants present in the source word. For example, *açúcar*, when loaned into Maxakalí, is [ã'tʃouɕ], with the surprising appearance of a nasal vowel. There are very few morphemes in Maxakalí that begin with oral vowels (e.g. negation *a-*, interrogative *ok-*), and only a handful of words to our knowledge, e.g. *onop* ‘bee’, that are not a suffix, onomatopoeic, or a loanword. What this overwhelming distribution of nasal vowels at the left edge of Maxakalí, even spontaneously so in loanwords, suggests, is that the ‘default’ specification at the left edge of Maxakalí words is nasality. This line of reasoning follows an observation in Rodrigues (1986) (further developed in Sândalo & Abaurre (2010) and Fujimura (2010)) that in a great deal of lowlands South American languages, the default resting position of the articulators beginning from silence is to have the velum lowered and to produce nasalization, unless otherwise interrupted by an explicit lexical [-nasal] specification. However, the phenomenon in itself remains surprising, and therefore is exactly a domain in which instrumental verification is of use. Consider, therefore, the following Intensity comparisons. The transcription accompanying the segmentation is in X-Sampa (see Appendix B for correspondences with the IPA). In all figures throughout this paper, nasal intensity is plotted with a solid line, and oral intensity with a dotted line.



This example was elicited with young female speaker EM; similar curves were found for the other speakers, and in this paper, we generally include one representative visualization per word token of interest, for reasons of space. As can be clearly seen in the diagram, the first vowel is nasal, with an equal intensity of oral (dotted) and nasal (solid) tracks, which extends into the initial portion of the affricate. Nasality does not reach this level any later within the word, and the second vowel is clearly more oral than the first. This example thus confirms a degree of nasalization on the initial vowel, in concord with previous descriptions.

A second loanword noted to exhibit this phenomenon results from the adaptation of *espelho* ‘mirror’, adapted into Maxakalí as [ʔĩjˈpæjç], shown below.

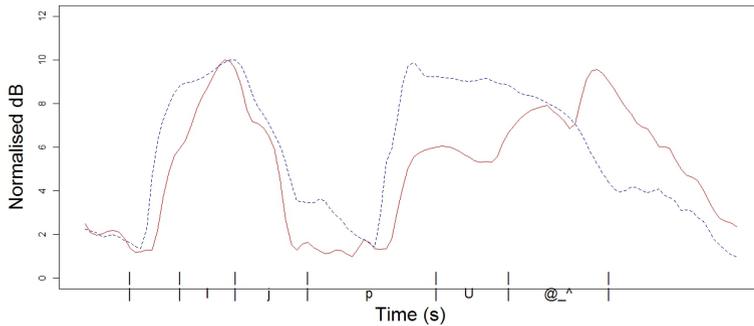
FIGURE 2 – [ʔĩjˈpæjç]



This example, from young male speaker IM, demonstrates a profile parallel to that of [ãˈʔou]: the first vowel shows an equal nasal and oral intensity, extending into the offglide, and then no longer present during the closure and release portion of the following stop, nor in the second vowel of the word. The intensity curves therefore again confirm the auditory transcription of spontaneous nasalization of the initial vowel.

In a third example, however, elicited with the loanword *espora* ‘spur’, its adaptation into Maxakalí did not show this initial nasalization [ɪjˈpʊə]. This token, recorded with adult male speaker GM, does not show nasalization.

FIGURE 3 – [ij'pʊə]



In this example, in contrast with the other two presented above, nasalization never reaches the same level of intensity as the oral profile in the initial vowel. We may thus conclude that initial nasalization did not take place with this token. There are a few possible explanations for its absence here. The first is the fact that this particular token refers to spurs (as found on a cowboy's boots), which may be a word of less familiarity and lesser usage to the speaker in question. The second is that this particular speaker has a higher degree of bilingualism, using Portuguese in his daily work. Comparison of the same token with older speakers, or of different tokens with the same speaker, could potentially arbitrate between these explanatory paths.

To conclude this subsection, we have demonstrated the utility of nasality measurements in bolstering the empirical support for the description of spontaneous initial nasalization. As such, the enterprise of continuing its theoretical explanation, in terms of a potential articulatory 'default' setting of left-edge nasalization, along the lines of Rodrigues (1986), may be further elaborated.

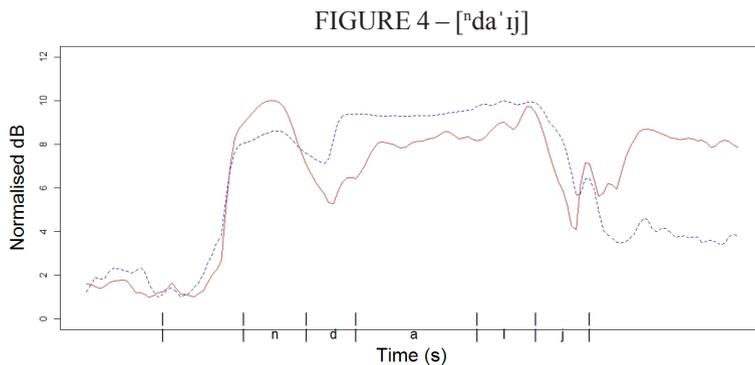
3.2 Prenasalization of stops

As has been documented throughout the literature on Maxakalí, from Gudschinsky, Popovich & Popovich (1970) onwards, the language has the presence of prenasalized stops such as [ʰd] word-initially. Prenasalization of voiced stops is a pervasive feature throughout lowlands South American languages, although with different potential underlying causes. For example, as discussed by Wetzels & Nevins

(2016), prenasalization may reflect ‘hypervoicing’, the mechanism discussed by Iverson & Salmons (1996) and Piñeros (2003), in which lowering the velum during the initial closure portion of a voiced stop allows maintenance of the pressure differential between subglottal and supraglottal regions that enables continued vibration of the vocal folds.

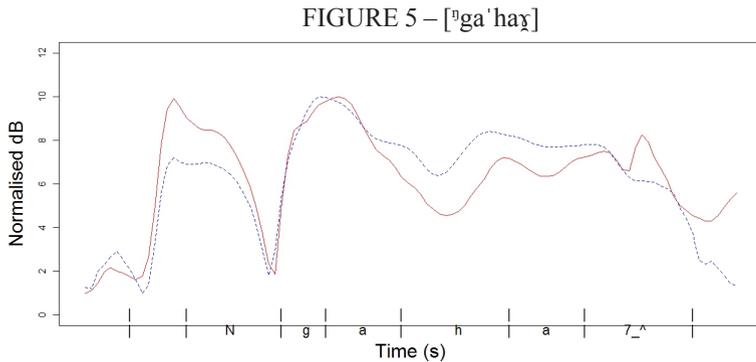
Differently from Southern Jê languages such as Kaingang, Maxakalí does not have oral-nasal contour segments [^hn] or circumoralized stops [^hn^h], which Wetzels & Nevins (2016) argue is the hallmark of a different underlying force, namely one to shield neighboring oral vowels from potential coarticulatory contamination by an adjacent nasal consonant. As such, Maxakalí’s prenasalized stops reflect the articulatory motivation of ‘venting’ an underlying voiced stop, and therefore documenting the existence of the nasal realization is an important aspect of the description to confirm. We include here three tokens: one is a Maxakalí noun, the second a loanword from Portuguese, and the third, a Maxakalí verb.

The first token presented here is [n^hda'ij], ‘clay pot’, as recorded with adult male speaker PM.



This example shows a clear prenasalization phase preceding the closure and release portion of the voiced stop that follows it. The velum remains raised for the remainder of the word until after the final consonant, at which point in the nasal track may simply reflect an outward breath and/or a return to the resting state of the articulators. From tokens such as this one, we can conclude that Maxakalí possesses prenasalization, the timing profile of which is easily enabled by the earbud recording method.

The second token chosen for inspection is the loanword [ˈgaˈhax̃], from the Portuguese *garrafa* ‘bottle’. The presence of prenasalization in this example bolsters the explanation of this phenomenon as reflecting hypervoicing of an underlying oral voiced stop, as the source word is clearly non-nasal.

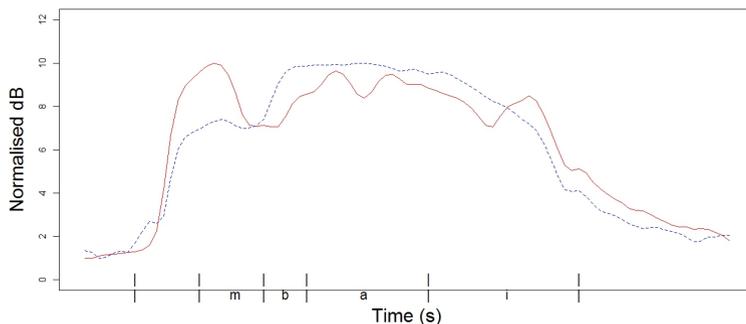


In this example, recorded with elderly female speaker MSM, there is a clear nasal portion, which in fact extends all the way into the beginning of the vowel. This speaker, therefore, has lowered the velum to achieve venting, and shows delayed closure to a purely oral airflow.

The final example of prenasalization considered here is with the verb [mbaˈi] ‘to be good’. Although this may be considered an adjective, as discussed by Campos (2009), it takes the same agreement prefixes as verbs. In fact, these agreement prefixes often inhibit (or bleed) the realization of prenasalization of the underlying voiced stop, as this prenasalization only reliably takes place in absolute word-initial position.² However, in this particular token, adult male speaker IZM does not produce it with the person prefix, which is common given that this particular lexical item can be used to mean something like “okay”, without a referential subject.

² Campos (2009, p. 18-19) contends that this prenasalization in verbs is one of the realizations of the third person prefix. However, if the prenasalization were a form of the third person prefix, it would be hard to explain why alienable nouns (which cannot occur with such prefixes) also show prenasalization in the same contexts.

FIGURE 6 – [ᵐba'i]

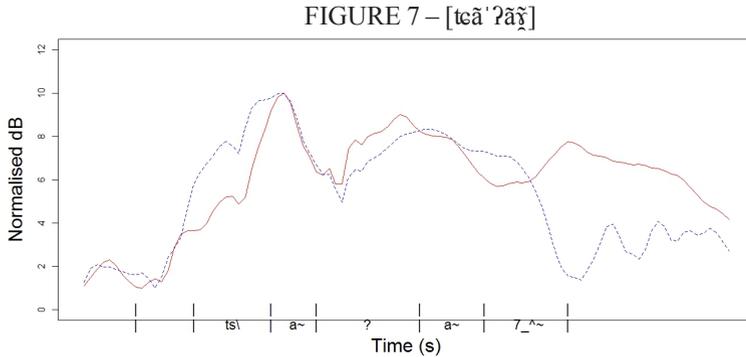


Much like the realization of [ᵐda'ij] above, this example shows a clear distinction between a prenasalization phase and an oral closure and release that follow it. While prenasalization is less often encountered in verbs than nouns for the reasons explained above, we conclude that it simply reflects affixation bleeding the otherwise expected process of nasal venting before a voiced stop, potentially applicable in both nouns (whether native or loanwords) and verbs.

3.3 Glottal transparency

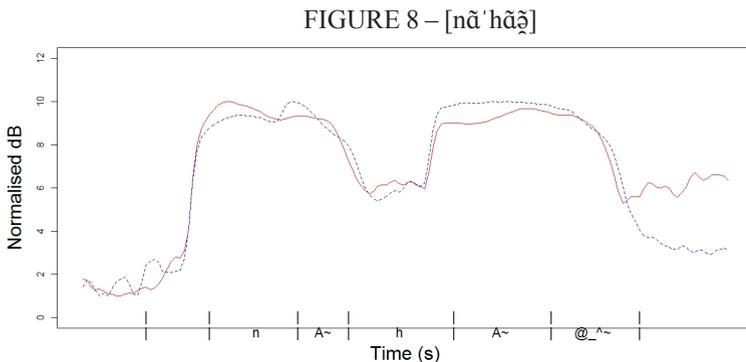
Maxakalí is a language that permits leftward nasal harmony when originating with a stressed nasal vowel, though this harmony is blocked by voiceless obstruents intervening, as is well-known within the typology of blocking in nasal harmony (WALKER, 2003). Interesting, therefore, is the question of whether the relevant blocking factor interacts with sonority and with supraglottal articulation. The phonological characterization of harmony therefore highly depends on whether the formalization of blocking vs transparency is in terms of pure definition as fricative or stop, or whether glottals are treated as more sonorous (as in LOMBARDI, 2001) – and hence permissive of nasalization, or simply irrelevant as they contain no supralaryngeal node (as in SAGEY 1986). Nonetheless, descriptive issues remain about the nasalization of glottals in Maxakalí, largely because of the difficulty of perceiving it auditorally. The use of nasal measurements is thus quite informative, particularly in regards to the contrast between the glottal fricative and the glottal stop.

The first word to be presented in this connection, the long form [tɛã'ʔãĩ] ‘slug’ is of importance because it demonstrates nasal transparency, originating with the stressed nasal vowel, and spreading leftwards across a glottal stop.



This token, produced by adult female speaker SM, shows that while there is no nasalization during the word-initial voiceless affricate, the velum is lowered immediately when the first vowel begins, and continues all the way to the end of the word, remaining steady throughout the glottal stop and onto the second vowel, and only decreasing during the final offglide. The presence of nasalization in this word – both in the glottal and indeed in the first vowel itself – is of importance in descriptive terms because this word has been previously described as lacking nasalization on the initial vowel (POPOVICH; POPOVICH, 2005).

The second token is one with the glottal fricative, [nã'hãĩ] ‘annato’, as produced by elderly female speaker MSM.



This example, like the one above, demonstrates nasality both in the first vowel and in the glottal fricative that intervenes between the first and second vowel. In fact, this word demonstrates nasality entirely throughout, providing a canonical example of a fully harmonized word. We may thus conclude that although they are distributionally more restricted in the language ([h] occurring only in onset, and the glottal stop occurring only intervocally), and while not phonologically contrastive in the language, glottal consonants in Maxakalí undergo nasalization in the same environments as voiced onsets.

The relevance of glottal transparency in Maxakalí takes on further importance given its interaction with the morphophonology of what are called long and short forms in the language. In its most compact description, long forms are found as the citation form of nouns, and found when they occur postverbally, e.g. [mã' hãḥ] 'fish'. However, in compounds, suffixed forms, and when occurring preverbally, such nouns occur in their short form, e.g. [mãḥ' piɥə] 'fish catch'. There is debate as to which form of the noun is underlying and which is derived; Araújo (2000), among others, argues that the long form is underlying, whereas Wetzels (2009)³ and Silva (2016) argues that the short form is underlying. One of Silva's arguments is based on the identity of the two vowels in the long form and their interaction with allophonic environments. In Maxakalí, vowels show a different realization when they precede coronal consonants: /a, ε, i, o/ → [ɑ, æ, ɪ, ʊ] / ____ [+coronal]. Importantly, the pre-coronal allophone is maintained in both halves of the long form, even though the 'first' half is not directly before a coronal. Thus, the long form of /dãd/ 'annatto', as shown above, maintains this allophonic vowel quality in both halves [nã' hãḥ], and not *[nã. 'hãḥ], demonstrating that the allophony is copied over to the first vowel in the long form. This pattern is more clearly explained if the short form is underlying. However, it may still be possible to explain this in terms of the long form being underlying, provided a mechanism for harmony (and thus vowel identity) is in place. Confirmation of the transparency of glottals

³ Although Wetzels' (2009) analysis considers that the short form is underlying, it relies in part on inaccurate data and affirmations from Araújo (2000) such as "[...] according to Araújo (2000), the glottal sound that occurs in the long forms, although it is usually realized as [h], alternates freely with [ʔ]." There are not any cases of such variation in our data.

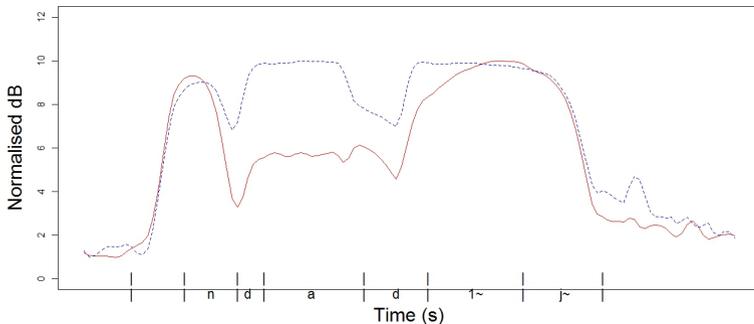
for nasal harmony, therefore, indirectly bears consequences for analyses of harmony more generally in the language.

3.4 Loanwords

The distribution of nasality in loanwords is of great interest, as discussed in Wetzels (2009) and Silva & Nevins (2015), particularly when it reflects a discrepancy in nasality of consonants between the source and donor language. In Maxakalí native vocabulary, syllables such as [na] and [dã] are banned: onset consonants cannot be nasal without the presence of a nasal vowel that follows. This predictable pairing of nasal onsets with nasal vowels has been reflected in an allophonic analysis of the two, whereby voiced stops are taken to be underlying and their nasal variant arises from predictable processes of spreading, and is enshrined in the orthographic system developed by Popovich & Popovich (2005), which does not graphemically distinguish voiced stops from their nasal counterparts. However, the influx of Portuguese loanwords is leading to the emergence of oral voiced stops before nasal vowels, as with words such as [bã'εɔ] (from Portuguese *banheiro* ‘bathroom’).

These divergences from the otherwise regular spreading of nasality within syllables found in Maxakalí are of course understandable in terms of the tradeoffs that arise in loanword phonology, whereby faithfulness to the identity of consonants in the source language play a role as well. Such a case arises in the adaptation of words such as *laranja* ‘orange’ in Portuguese, which are adapted as [ˈda.dĩj]. As Maxakalí has no liquids in its native inventory (despite the name of the language, which remains an exonym), both the lateral and rhotic are adapted as [d]. Notably, this coronal onset resists onset nasalization.

FIGURE 9 – [ˈda'dĩj]



The recording, with adult male speaker RM exhibits prenasalization of the word-initial [d], of the same nature as that described in the subsection above. Of central interest to the present discussion is the predominately oral onset in the second syllable. While the final vowel and its coda are fully nasal, as shown by the solid intensity curve, its onset fails to undergo nasalization. This example therefore provides confirmation for a fully oral onset consonant preceding a nasal rime in Maxakalí, arguably an emerging consequence of the sometimes conflicting demands posed by loanword phonology.

As described in Wetzels (2009) and Silva & Nevins (2015), on the other hand, occasionally loanwords do exhibit a distribution of nasality that goes far above and beyond what is present in the source form. The adaptation of the Brazilian soccer team *Flamengo* is of particular interest in this case. This item undergoes epenthesis between the initial labial consonant (adapted as voiceless [p] in Maxakalí) and the following coronal consonant. However, speakers show three possibilities of the realization of the source [l]. Some adapt the lateral as [n] as in [pãñãmãẽ], undergoing nasal spreading from the following nasal vowel. However, these fall into two subgroups: [pãñãmãẽ] and [panãmãẽ] where the difference between these two lies in the presence (the former) or absence (the latter) of further leftwards vowel harmony extending to the preceding vowel. These are shown below for speakers RM and IZM below:

FIGURE 10 – [pãñãmãẽ]

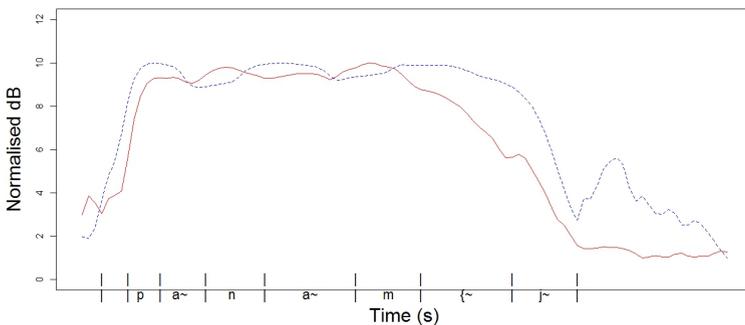
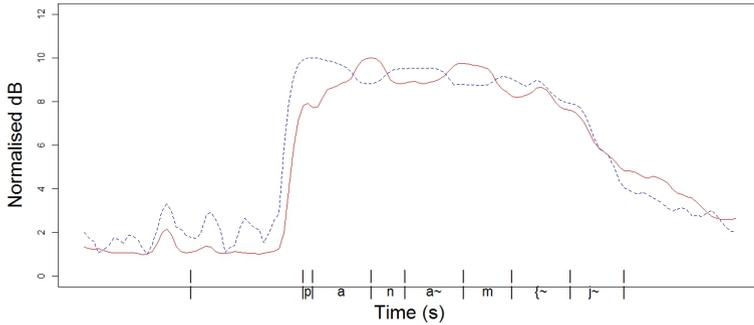


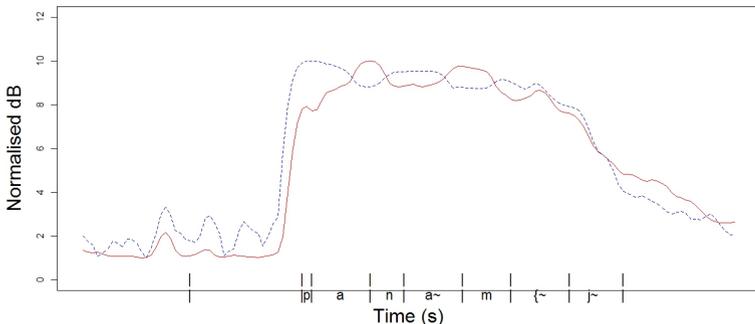
FIGURE 11 – [panã'mãẽ]



As can be seen from the solid intensity curves above, adult male speaker RM shows nasality from the first vowel of the word all the way until the end, with a gradual raising of the velum as nasality peters out. On the other hand, adult male speaker IZM only begins nasality in the final portion of the initial vowel, as coarticulatory anticipation of the following nasal onset, after which it extends all the way to the end of the word. From this instrumentally-informed comparison, we can conclude, in concert with the proposal of Silva & Nevins (2015), based on purely auditory transcriptions that leftwards vowel-to-vowel harmony is a gradient phenomenon, even across an intervening nasal consonant.

Still others, in a solution to loanword phonology that remains within the native inventory (in eschewing [l]) but diverges from its phonotactics, follow the pattern of [ˈda.dĩ] as [padãmãẽ]. This is shown below for the young male speaker IM.

FIGURE 12 – [padã'mãẽ]



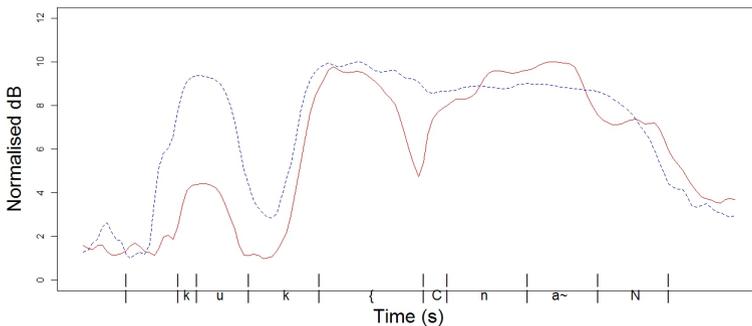
In this example, as shown by the solid intensity curve, velum lowering only occurs during the nasal vowel [ã] following the oral stop [d], and continues until the end as in the same relevant portion of the profile of this word as pronounced by RM above.

Summarizing this subsection, both words discussed here illustrate that Maxakalí loanwords allow an oral voiced stop to precede a nasal vowel, in contradiction to the otherwise entirely predictable allophonic distribution whereby nasal vowels must have nasal onsets (if the consonant is voiced). Given that such a pattern necessitates a revision of the overall picture of Maxakalí phonology (and potentially, down the road, its orthography), this makes all the more valuable a verification that indeed such consonants evade nasality, afforded by this instrumental method.

3.5 Nasal spreading from the diminutive

In this final subsection, we discuss one of the more language-specific aspects of the distribution of nasality in this language, one resulting from the leftward spreading induced by the diminutive suffix [-nãŋ]. According to Campos (2009), this suffix induces nasal spreading in verbs, but not nouns. This is shown for [kukæç'nãŋ] ‘dog-dim.’, by the adult male speaker PM, below.

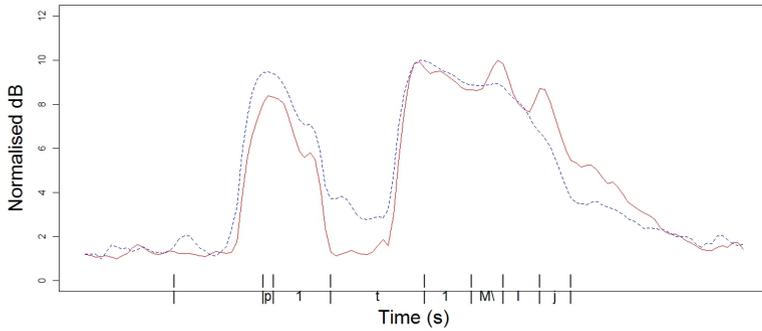
FIGURE 13 – [kukæç'nãŋ]



As shown by the red intensity curve, nasality is not part of the production of the root, and obtains only during the suffix-initial nasal consonant, thereby demonstrating that, for nouns such as this one, the diminutive does not induce nasalization.

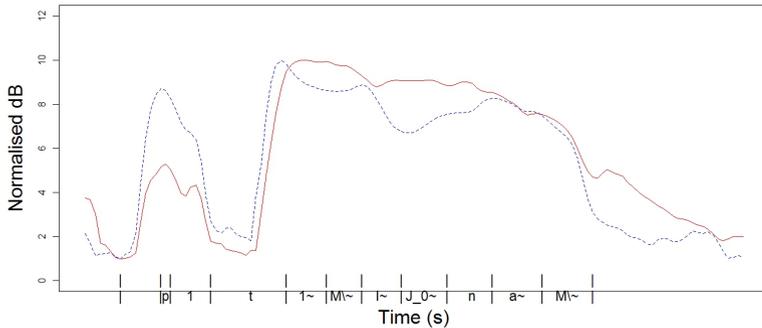
In verbs, however, as discussed by Campos (2009), the diminutive causes leftward nasalization of the root-final consonant and, in turn, the vowel that precedes it. This can be shown in the following pair of recordings with adult male speaker IZM below. The first curve is for [pi'tiuj] 'heavy', which has no nasality. The second is for its diminutivized form 'heavy-dim.', which shows nasalization of the entire rime of the root.

FIGURE 14 – [pi'tiuj]



In the figure above, the nasal curve does not surpass that of the oral curve. By contrast, in the figure below for [pitũĩĩj' nãũ] recorded with IZM as well, nasality clearly obtains from the root-final vowel onwards:

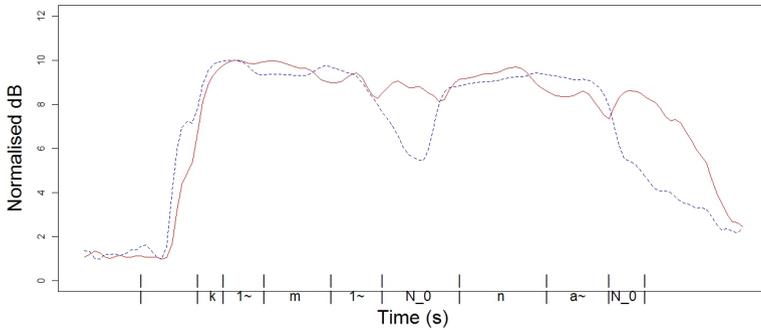
FIGURE 15 – [pitũĩĩj' nãũ]



This latter example, ‘heavy-dim.’, based on a verb, shows a clear effect of morphologically-induced nasalization into the base. Its morphological conditioning is evident in the fact that in a nearly parallel noun, [pʰitʰiɰjˈnãũ] ‘bird/birdy’, no nasalization obtains. Of course, this is a ‘frozen’ diminutive, based on a root that is no longer synchronically present,⁴ but it also serves to show that the segmental content of the final syllable [nãŋ] is not enough on its own to induce nasalization in an otherwise identical sequence to its left.

Finally, it should be said that diminutive nasal spreading in verbs itself shows interspeaker variation. The form [kʰiˈbʰik] ‘bad’ can undergo complete nasalization, as shown for speaker IM below:

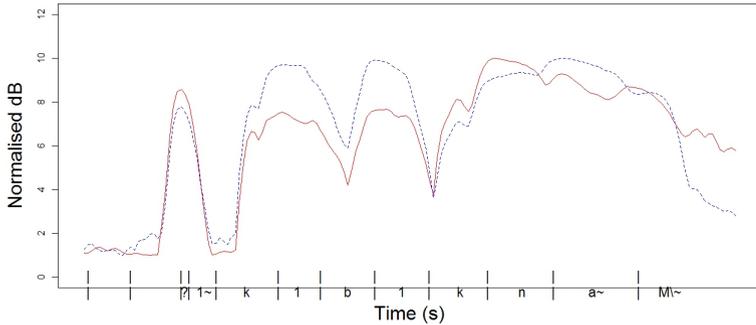
FIGURE 16 – [kʰimĩjˈnãŋ]



This form, given the base [kʰiˈbʰik] ‘bad’, has undergone leftward nasalization in the root-final coda, subsequently spreading to the entire syllable, and further leftward in nasal harmony, halted only by the word-initial voiceless stop. On the other hand, speaker MSM shows no nasalization at all of the root. Note that this token includes an inflectional 3rd person prefix [ĩ], which is an inherently nasal vowel.

⁴ The plausibility of a root meaning ‘bird’ inside a frozen diminutive, interestingly, gains support from the recent tentative reconstruction of ‘feather’ and ‘bird’ in Proto-Macro-Jê as *prɣ(C) by Nikulin (2015, p.293), which would then undergo sound change to Maxakali *[pʰitʰj], following the steps posited in Silva (2016).

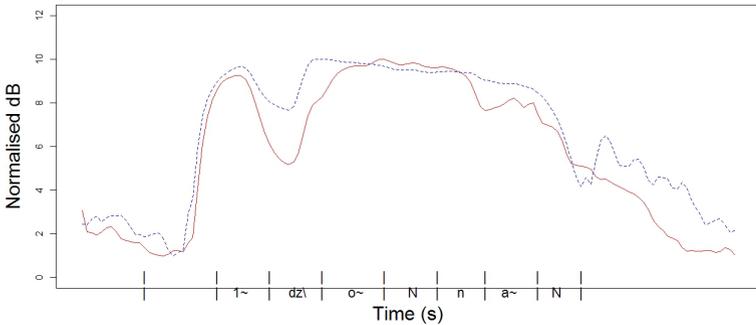
FIGURE 17 – [ĩkibik' nãũŋ]



In this token, the oral intensity curve dominates the nasal out the root, but the velum does lower during the production of the suffix.

Finally, we show that the pattern of nasalization induced by the diminutive suffix with monosyllabic roots such as [dʒouk] ‘straight’ can lead to a sequence of oral voiced stop followed by nasal vowel, again in apparent contradiction of the otherwise exceptionless phonotactic that nasal vowels induce onset nasalization of voiced stops.

FIGURE 18 – [ĩdʒõŋ' nãŋ]



The gradient patterning of nasalization found with verbs but not nouns, and with exceptional and optional behavior, can therefore lead to another source for emerging non-allophonic distribution of voiced stops vs. their nasal counterparts.

To summarize the empirical situation overall, therefore, nasalization in verbs, while variable in its application and extent, can

obtain, as confirmed by earbud microphony, while nasalization in nouns was never found. This is a valuable new descriptive contribution that goes beyond impressionistic descriptions of nasalization. The availability of earbud microphony for field recordings enables us to study phenomena such as this with a wide range of speakers, across ages and genders, which allows one to investigate variable phenomenon with a large sample size and address ethnosociolinguistic questions. Recordings with a diverse demographic range of participants would not have been as straightforward if one needed to bring speakers far from their homes to a laboratory and/or use unfamiliar equipment.

4 Future directions

In this article, we have shown how instrumental measurements of nasal airflow can break the stalemate that can arise with descriptive claims that necessitate potential revisions from one's theoretical comfort zone. With earbud microphones, segmental transcription, and subsequent execution of scripts, questions such as prenasalization of voiced stops, spontaneous nasalization of initial vowels, failure of onset nasalization, laryngeal transparency to nasal harmony, and morphophonologically-conditioned restrictions on nasal spreading can all be addressed with visual clarity.

As these five aspects of nasalization across consonants, vowels, and their interaction with processes such as prenasalization, nasal harmony, and loanword adaptation are widespread across South American languages (and beyond), the promise of this technique for such crosslinguistic investigations is clear. In some cases, one can confirm earlier auditory impressions, in others apparently divergent descriptions or analyses do turn out to boil down to interspeaker variation, and in others, the description or the researcher can reflect on potential reasons the auditory impressions diverge from the representations obtained by recording and normalizing oral and nasal tracks. Either way, more data and metadata accompanying them are added to the empirical base.

Going beyond these considerations, we also view the potential of earbud microphony and its nearly immediate ability to generate a visual representation of the nasal profile of a word as yielding possibilities for experimental linguistics to show increased integration with language pedagogy. Showing native speakers and in particular, those concerned with indigenous education and orthography reform the immediate

representation of nasality has clear potential for more close integration between educational technology and development of pedagogical materials within the village context. As an example, we can cite the fact that many Maxakalí often omit the orthographic tilde when writing. This diacritic serves as an indicator of nasalization and its absence can lead to potential confusion in reading. Its omission – which could be influenced by the fact that is impossible to place tilde over the graphemes *e,u,i* using a Portuguese keyboard layout, and so users must use <ê,û,î>, the comfortability of typing which may lead to questioning how often they are needed – may be because some instances are highly predictable thanks to nasal harmony from the stressed vowel or by other means.

Nasal earbuds are easy to use and lend themselves to excellent possibilities for student projects at the beginning research level as well as for younger or technologically-savvy members of the indigenous community to be involved hands-on with new ways of visualizing the sound patterns of their language. We have shown that they can be used for consonantal and vocalic nasality phenomena, loanword phonology, variable phenomena, and morphologically-conditioned patterns, and in making our dataset available with the textgrids and a pointer to the scripts, they have the potential to form part of ‘visual dictionaries’ showing what nasality does within given sets of recurrent word patterns.

Acknowledgements

We would like to thank Martin Kohlberger and Jesse Stewart for help with numerous technical aspects of applying this research method. We would also like to thank Leo Wetzels and the three anonymous reviewers for their thorough reading and insightful suggestions that led to improving this paper. Finally, we acknowledge the Maxakalí people from Aldeia Verde for their contribution and collaboration.

References

ARAÚJO, Gabriel Antunes. *Fonologia e morfologia da língua Maxakalí*. 2000. Thesis (Masters’) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2000.

AUDIBERT, Nicolas; AMELOT, Angélique. Comparison of nasalance measurements from accelerometers and microphones and preliminary development of novel features. In: INTERSPEECH 2011 - ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION, 12th, 2011, Florence. *Proceedings...* Florence: International Speech Communication Association, 2011. p. 2825-2828.

CAMPOS, Carlo Sandro. *Morfofonêmica e Morfossintaxe do Maxakalí*. 2009. Dissertation (Doctoral) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

CHEN, Marilyn. *Acoustic Correlates of Nasality in Speech*. 1996. Dissertation (PhD) – MIT, Cambridge, MA, 1996.

DEMOLIN, D. The influence of aerodynamic constraints on the shape and the dynamics of phonological systems. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, XVII., 2011, Hong Kong. *Proceedings...* Hong Kong: City University of Hong Kong, 2011. p. 60-63.

FLETCHER, Samuel G.; SOOUDI, IRADJ; FROST, Sharon D. Quantitative and graphic analysis of prosthetic treatment for ‘nasalance’ in speech. *The Journal of Prosthetic Dentistry*, Elsevier, v. 32, p. 284-291, 1974.

FUJIMURA, Osamu. *Syllable structure of Kaingang: A C/D Model Interpretation*. Presented at TCP, July 2010.

GUDSCHINSKY, Sarah; POPOVICH, Harold; POPOVICH, Frances. Native Reaction and Phonetic Similarity in Maxakalí Phonology. *Language*, Washington, v. 46, p. 77-88, 1970.

IVERSON, G. K.; SALMONS, J. C. Mixtec prenasalization as hypervoicing. *International Journal of American Linguistics*, Chicago University Press, v. 62, p. 165-175, 1996.

LOMBARDI, Linda. Why place and voice are different: constraint-specific alternations in Optimality Theory. In: _____. (Ed.). *Segmental Phonology in Optimality Theory*. Cambridge University Press, 2001. p. 13-45.

MONTAGU, Julie; AMELOT, Angélique. Comparaison des apports de différents méthodes d'enregistrement de la nasalité: étude préliminaire. In: RENCONTRE JEUNES CHERCHEURS, 2005, Toulouse. Toulouse: Université Paul Sabatier, 2005. p. 17-21.

NIKULIN, Andrey. Apofonia e o sistema vocálico do Proto-Jê Meridional: contribuição para estudos comparativos das línguas Jê. *Revista Brasileira de Linguística Antropológica*, Brasília, UnB, v. 7, n. 2, p. 275-306, 2015.

PIÑEROS, Carlos Eduardo. Accounting for the instability of Palenquero voiced stops. *Língua*, Elsevier, v. 113, p. 1185-1222, 2003.

POPOVICH, H.; POPOVICH, F. *Dicionário Maxacali-Português com Glossário Português-Maxacali*. Cuiabá: SIL, 2005.

RODRIGUES, Aryon. Nasalização e fronteira de palavra em Maxakalí. In: ENCONTRO NACIONAL DE LINGUISTICA, V., 1981, Rio de Janeiro. *Anais...* Rio de Janeiro: Pontifícia Universidade Católica, 1981. p. 305-311.

RODRIGUES, Aryon. Silêncio, pausa, e nasalização. In: ENCONTRO NACIONAL DE LINGUISTICA, VIII., 1986, Rio de Janeiro. *Anais...* Rio de Janeiro: Pontifícia Universidade Católica, 1986. p. 153-158.

SAGEY, E. *The representation of features and relations in non-linear phonology*. 1986. Dissertation (Doctoral) - MIT, 1986.

SÂNDALO, F.; ABAURRE, M. B. Orality spreading in Pirahã. *Liames*, Campinas, Unicamp, v. 10, p. 7-19, 2010.

SHOSTED, R. Vowel context as a condition for nasal coda emergence: aerodynamic evidence. *Journal of the International Phonetic Association*, Cambridge University Press, v. 36, p. 39-58, 2006.

SILVA, Mário André Coelho. *A coda consonantal em Maxakalí*. 2015. Thesis (Masters') - Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 2015.

SILVA, Mário André Coelho. *Um novo olhar sobre formas longas e curtas em nomes no Maxakalí*. In: ENCONTRO MACRO-JÊ, VIII., 2016, Londrina. Paper presented.

SILVA, Mário André Coelho; NEVINS, Andrew Ira. *Domain-based preservation of nasality in Maxakalí loanwords*. In: MANCHESTER PHONOLOGY MEETING (MFM), 23rd, 2015, Manchester. Paper presented.

STEWART, J.; KOHLBERGER, M. Earbuds: A method for measuring nasality in the field. *Journal of Language Documentation and Conservation*, University of Hawai'i, v. 11, p. 49-80, 2017.

STYLER, Will. *On the Acoustical and Perceptual Features of Vowel Nasality*. 2015. Dissertation (PhD) - University of Colorado, Boulder, 2015

WALKER, Rachel. Reinterpreting transparency in nasal harmony. In: VAN DE WEIJER J.; VAN HEUVEN, V.; VAN DER HULST H. (Ed.). *The Phonological Spectrum Vol 1: Segmental Structure*. Amsterdam; Philadelphia: John Benjamins Publishing, 2003. p.37-72.

WETZELS, Leo. Prevowels in Maxacali: Where They Come From. *Boletim da Associação Brasileira de Linguística*, Belo Horizonte, v. 14, p. 39-63, 1993.

WETZELS, Leo. Nasal Harmony and the Representation of Nasality in Maxacali: Evidence from Portuguese Loans. In: CALABRESE, A.; WETZELS, W. L. (Ed.). *Loan Phonology*. Amsterdam: John Benjamins Publishing, 2009. p. 241-270.

WETZELS, Leo; NEVINS, Andrew Ira. *Prenasalized vs Postoralized Stops: Diverse Types of Enhancement*. Manuscript, 2016.

WETZELS, Leo; SLUYTERS, Willebrord. Formação de raiz, formação de glide e 'decrowding' fonético em Maxacalí. In: WETZELS, L. (Ed.). *Estudos Fonológicos das Línguas Indígenas Brasileiras*. Rio de Janeiro: Editora UFRJ, 1995. p. 103-149

APPENDIX A

Recorded Wordlist and Translation

Maxakalí	English
Ānenep	Andrew
āxok	sugar (< Port. açúcar)
āyuhuk	non-indigenous
gahap	bottle (< Port. garrafa)
hitūmnāg	happy-DIM
hitup	happy, healthy
Homet	Roberto (male name)
Hoyet	Rogério (male name)
Ĩã	Ian
ixōg ~ ĩxōg	bird
ixpot	spur (< Port. espora)
ĩypex	mirror (< Port. espelho)
kokexnāg	dog-DIM
Kōyāyiy	Cunhadinho (male nickname)
kũĩn	stripped
kũĩnnāg	stripped-DIM
kumuk	bad
kumuknāg ~ kũmũgnāg	bad-DIM
kũnōhōn	cockroach
māhām	fish
mahap	turn around
mai	good
maināg ~ mǎĩnnāg	good-DIM
Māxakani	Maxakalí (exonym)
mayix	bowl (< Port. vasilha)
mīkax hã	with the knife (lit. knife INSTR)
mĩnkup	sugarcane
mot	ball (< Port. bola)
mot mōyōn	play football

mox	cow, ox (< Port. boi)
mūnūytut	cow, ox
nagnāg	dry-DIM
nāhān	annatto
nak	dry
Nanañn	Ladainha (town name)
nanuy	orange (< Port. laranja)
nax ~ naix	pot
Pānāmēy ~ Panāmēy	Flamengo (football team)
penāhā	see, watch
ponognāg	white-DIM
ponok	white
putōōy	mud, clay
putux	heavy
putuxnāg	bird
putūynāg	heavy-DIM
Tēān	Ateanderson (male name)
Tikmūūn	Maxakalí (endonym)
tonok	step
tonoknāg	step-DIM
xāām	snail
xēēnāg	true
xuta	red
xutānnāg	red-DIM
Yaet	Isael (male name)
yīmamnāg	turn around-DIM
yīmap	turn around
Yitmā	Gilmar (male name)
yognāg	straight-DIM
yok	straight

APPENDIX B

X-SAMPA values and IPA correspondence

X-SAMPA	IPA
i	i
I	ɪ
{	æ
a	a
A	ɑ
o	o
u	u
U	ʊ
7_^	ɣ
@_^	ʁ
p	p
t	t
ts\	tʃ
k	k
?	ʔ
b	b
d	d
dz\	dʒ
g	g
m	m
n	n
J	ɲ
N	ŋ
h	h
C	ç
j	j
M\	ɥ
~ (e.g.: a~)	nasality (e.g.: ã)
_0 (e.g.: n_0)	devoicing (e.g.: ŋ)

Social affective variations in Brazilian Portuguese: a perceptual and acoustic analysis

Variações socioexpressivas no português do Brasil: análise perceptiva e acústica

Albert Rilliard

Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI), Centre National de la Recherche Scientifique (CNRS), Université Paris-Saclay, Orsay / France

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
CNPq

albert.rilliard@limsi.fr

João Antônio de Moraes

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
CNPq

jamoraes3@gmail.com

Resumo: Este trabalho descreve e analisa um *corpus* de 17 atitudes prosódicas produzido por 21 falantes do português do Brasil (PB). A interpretação dada a essas atitudes baseia-se em um experimento de escolha livre (*free-labelling*) do qual participaram 22 ouvintes brasileiros. O agrupamento das diferentes expressões obtidas no experimento é relacionado aos parâmetros acústicos da prosódia (F_0 , intensidade, duração), e sua interpretação é feita com base na literatura sobre a manifestação da expressividade na fala e nas teorias do uso simbólico da voz em interações (código da frequência, de Ohala, e código do esforço, de Gussenhoven). Os resultados mostram que as principais dimensões de sentido identificadas pelos ouvintes podem ser reagrupadas em três grandes categorias: valência positiva, expressões assertivas e expressões

dubitativas, dimensões que se correlacionam com as medidas acústicas.

Palavras-chave: prosódia; *free-labelling*; dimensões de sentido; medidas acústicas.

Abstract: This work presents a description and an analysis of an attitudinal corpus recorded in Brazilian Portuguese by 21 speakers. The interpretation of their performances across 17 situations, by L1 speakers of Brazilian Portuguese, is studied thanks to a free-labelling experiment. The grouping of expressions is then related to acoustic parameters of prosody. The interpretation of these descriptions is done in the light of literature on expression of affect in voice, and theories of symbolic use of voice in spoken interactions (Ohala's Frequency Code, and Gussenhoven's Effort Code). Results show that the main dimensions of meaning retrieved by the listeners may be summarized as: positive valence, assertive and dubitative expressions. These dimensions do correlate with the acoustical measures.

Keywords: prosody; free-labelling; dimensions of meaning; acoustic measures.

Recebido em: 5 de janeiro de 2017.

Aprovado em: 14 de março de 2017.

1 Introduction

Communication, during face-to-face interactions, uses a multimodal framework of signals to accurately convey the targeted speech acts. This includes gestures (WU; COULSON, 2007), facial expressions (GONZÁLEZ-FUENTE *et al.*, 2015), lexical choices and prosodic variations (BOROD *et al.*, 2000). Use of prosodic cues has received a sustained attention. Researches in various languages have studied its uses to signal various speech acts, or sets of speech acts: e.g. in Chinese (CHANG, 1958), English (ULDALL, 1960), French (FÓNAGY; BÉRARD, 1972), Japanese (FUJISAKI; HIROSE, 1993). Works on prosodic attitudes focus on one specific language, listing the various prosodic changes observed in speech, often for foreign language teaching purposes (e.g. DELATTRE, 1963; MARTINS-BALTARD,

1977). These approaches end with lists of labels describing the meaning of the described prosodic changes. Such lists raises problems as soon as one tries to compare them cross-culturally: the folk labels (i.e. labels used with their vernacular meanings) fall under Wierzbicka's criticism on the notional variation lying under folk labels' translations—that do not bear equivalent concepts (WIERZBICKA, 1985). Thus the concepts behind e.g. the term “*irony*” in English shall not be exactly comparable to those evoked by “*ironia*” in Portuguese.

To study social affects expressed by prosodic means, and to compare them cross-culturally (both in production and perception), a paradigm as been set up (RILLIARD *et al.*, 2013) that proposes to speakers of various languages to produce the same simple sentence, in various contexts (cf. FÓNAGY; BÉRARD, 1972, for a similar practice), thus eliciting a range of vocal expressions. The same contexts, defining the speaker's communication goals, her/his relationship to the interlocutor (in terms of social proximity and hierarchy: SPENCER-OATEY, 1996), introduced short dialogues that end with the targeted prosodic expression.

This paper presents a description of the recordings process in Brazilian Portuguese (BP), acoustical analyses and perceptual interpretations of the prosodic performances, which follow these principles. The findings are discussed in the light of similar results in other languages—and together with interpretations linked to theoretical account of intonational meaning (OHALA, 1994; BANSE; SCHERER, 1996; GUSSENHOVEN, 2004; GOUDBEEK; SCHERER, 2010).

2 Corpus and method

2.1 Capturing cross-culturally comparable social affects

This study is part of a project aiming at comparison of prosodic performances between languages and cultures (RILLIARD *et al.*, 2013). To that aim, and to avoid the bias possibly introduced by the use of folk labels, speakers were set up in situations where they had to interact with an interviewer. Small scenarios were written that end with the speakers uttering a sentence with a speech act and an attitude defined by the scenario—e.g. asking politely for a fruit, or requesting this fruit with authority. These scenarios allow the production of two target sentences (for the BP version: sentence B “*uma banana*”: a

banana, and sentence M “*Maria dançava*”: Maria was dancing) with 17 different speech acts and/or attitudes (note for BP, a 17th scenario was added to produce “*descrédito*”, discredit). The use of scenarios allows expressive variations that convey similar communicative values without resorting to folk labels. The scenarios were originally written in English, and then translated in various languages (French, Japanese, German, Cantonese, and of course BP) to enable the production of the same attitudes in each of these languages. The details of the 17 situations used for introducing the performances of the sentence “*Maria dançava*”, with the corresponding dialogues, are presented in annex. We give here the names of the targeted expressions, with their English translations and abbreviations: “*declarativo neutro*” (neutral declarative sentence, DECL); “*pergunta neutra*” (neutral question, QUES); “*admiração*” (admiration, ADMI); “*arrogância*” (arrogance, ARRO); “*autoridade*” (authority, AUTH); “*desprezo*” (contempt, CONT); “*evidência*” (obviousness, OBVI); “*incerteza*” (uncertainty, UNCE); “*ironia*” (irony, IRON); “*irritação*” (irritation, IRRI); “*pergunta com estranheza*” (doubt, DOUB); “*pisando em ovos*” (walking on eggs, WOEG); “*polidez*” (politeness, POLI); “*sedução*” (seduction, SEDU); “*sinceridade*” (sincerity, SINC); “*surpresa*” (surprise, SURP); and “*descrédito*” (discredit, DESC).

The situation coined as “walking-on-eggs” in English corresponds to an adaptation of what is called *kyoshuku* in Japanese: a concept that does not have an adequate translation in English or Portuguese, and is described by Sadanobu as “corresponding to a mixture of suffering ashamedness and embarrassment, which comes from the speaker’s consciousness of the fact his/her utterance of request imposes a burden to the hearer” (SADANOBU, 2004, p. 34). In order to study the prosodic performances expressing such a *kyoshuku* expression, as well as for the other expressions, speakers from various cultural origins were asked to interact in a communication context corresponding to the expressions of these social affects.

2.2 Recordings and acoustic measures

Twenty-one speakers (10 females) having Brazilian Portuguese, in the variety of Rio de Janeiro, as their L1 were recorded. Recordings took place in a sound-treated room at the *Laboratório de Fonética* of UFRJ, Rio de Janeiro. Each speaker had first recorded the corpus in one

of her/his L2 language (in this case, either Japanese, English, or French), then record the corpus in BP, each time interacting with a L1 speaker of the target language. The present paper focuses on the recordings in BP.

The speakers were seated in front of a Panasonic AG-AC160 video camera, equipped with an Earthworks QTC1 omnidirectional microphone. The microphone was placed at one meter of the speaker's mouth so to reduce intensity changes linked to body movements. The microphone level was calibrated before each recording session using a Brüel & Kjør acoustical calibrator: recordings were latter corrected for variation in the input level on the basis of this input sound. The two target sentences, “*Uma banana*” and “*Maria dançava*”, were extracted from the recordings, and then segmented at the level of phonemes using the Praat software (BOERSMA; WEENINK, 2016) and the *easyalign* plugin (GOLDMAN, 2011), with manual corrections of the automatic segmentations.

An acoustic analysis of the signal was done, extracting the mean fundamental frequency on each vowel (F_0 , expressed in semitones relative to 1 Hz, and measured using Praat default algorithm with hand correction for octave jumps and other errors), the syllabic duration (express in second), and the mean intensity on vowels (A-weighted intensity, expressed in dB-A: cf. LIENARD; BARRAS, 2013).

These acoustic parameters have been selected as reflecting the main dimensions of prosodic changes. Ohala (1994) proposes that the use of F_0 in vocal communication is mostly of symbolic origin: low / descending F_0 being related to assertive and dominant behaviours (reflecting the large size of the speaker), and high / rising F_0 being related to submissive and interrogative behaviours (linked to the smaller size of the speaker). Gussenhoven (2004) built on Ohala's “Frequency Code”, and proposes the existence of an “Effort Code”, that is related to the vocal effort exerted by the speaker while producing speech, and that also have influences on the voice's F_0 . Interpretations of this Effort Code link the involvement of the speaker in her/his spoken utterance (cf. DANEŠ, 1994 for this notion of involvement) and to the arousal of the expression (SCHERER, 2009; GOUDBEEK; SCHERER, 2010). On a speech production point of view, the vocal effort is related to the muscular tension at the glottis, and Titze & Sundberg (1992) as well as Traunmüller & Eriksson (2000) linked the production of variation of effort to the voice's intensity, and show there is a relation between

higher efforts and higher F_0 . Let's note that intensity levels were strictly controlled during the recording of this corpus, and avoid the use of e.g. spectral slope (HANSON, 1997) as a mean to retrieve a voice's vocal effort; this is fortunate given the high influence of vowel on such type of spectral measures, and their sensitivity to relatively small (and phonetically unbalanced) data. The three parameters have been standardized for intrinsic difference linked to speakers, so to keep only expressive variation—they are thus expressed in z-scores in this paper.

2.3 Perceptual analysis

In order to evaluate the expressive content of the recorded stimuli, they were submitted to a perceptual analysis. First, a performance test was run, asking listeners to judge the adequacy of each recording regarding to the targeted speech act/attitude. Listeners were 10 L1 speakers of BP trained in phonetics. They were presented with the 17 situations, and they observed the performances of speakers, knowing the targeted attitude. They had to rate the quality of the 21 speakers performances on a 1 (very poor) to 9 (excellent) scale. The scores given by each listener were standardized to remove variations in the individual use of the scale, and these z-scores were used to select the best performances in each of the 17 attitudes. The six best performers (3 females, 3 males) for each attitude and each sentence were selected. Among these 6 speakers, four (2 females, 2 males) were retained for a second perceptual analysis (the selection was made by the authors of this article). 20 speakers out of the 21 were selected for at least once of their performance through this process; the speaker who was not selected was not the one receiving the lowest overall score, but one that do receive medium rating for all attitudes. This sub-selection was done to ease the task of subjects in the second test, which cannot be performed on the 714 stimuli.

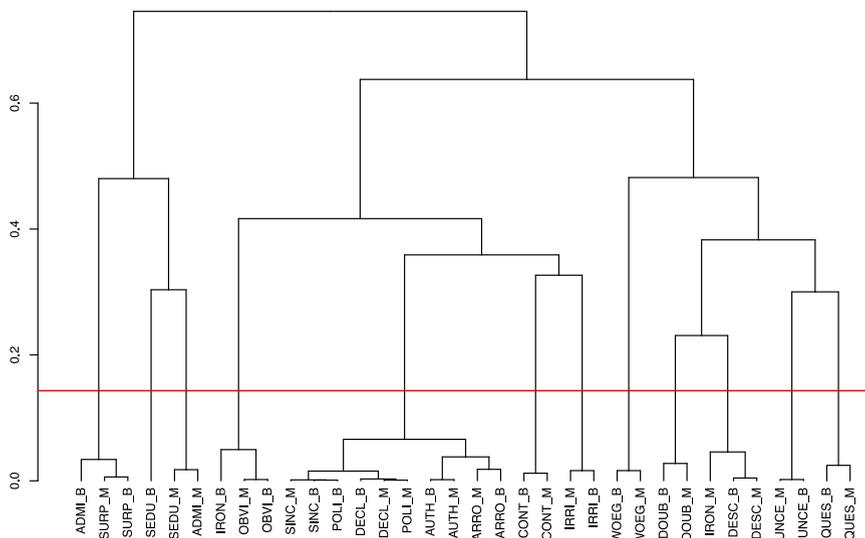
The second perception test was based on a free-labelling paradigm (cf. WIDEN; RUSSELL, 2003; GREENBERG *et al.*, 2009): the listeners, L1 speakers of PB, were asked to describe the expressivity of the presented audio-visual stimuli, using one substantive or adjective they think best describes the expression. They were presented with 136 stimuli (the two sentences in each of the 17 attitudes as performed by 4 speakers) randomly sorted for each listener. A typical experiment lasted 40 minutes. 22 listeners took part in the experiment (16 females; mean

age: 25 year-old). Listeners did not know about the recording procedure nor were instructed about the situations used to record the expression they were presented with. 530 different labels were provided by these 22 subjects, a number that was reduced to 274 after normalization: first typos were corrected; then, in case subjects type more than one word, the first was selected (as they were instructed to give only one), then marks of gender and plural were removed. On the remaining list, labels appearing in their adjective or verbal forms were converted to substantive when it was perceived to convey the same meaning (e.g. “*afirmação*, *afirmando*, *afirmar*, *afirmativa*, *afirmativo*” was encoded as “*afirmação*”, while “*inexpressivo*”, or “*sonhador*” were kept—as “*inexpressividade*”, or “*sonho*” are less frequent or have a different meaning). The frequency of each of these 274 labels, for the 17 attitudes in each sentence (i.e. 34 situations performed by 4 speakers each), was calculated and forms a large 34 by 274 matrix that was submitted to a correspondence analysis (using the FactoMineR library of the R software; HUSSON *et al.*, 2011; R CORE TEAM, 2016). The correspondence analysis aims at reducing the dimension of this large matrix, so to extract abstract dimensions that carry most of the information relative to the description of each type of stimulus by means of the labels. The analytic procedure is inspired from works on dimensions of meanings—typically those pioneered by Osgood and colleagues (OSGOOD *et al.*, 1957; OSGOOD *et al.*, 1975) using list of dimensions based on opposed labels (a procedure already used to study the dimensions of attitudinal meaning: ULDALL, 1960; FÓNAGY; BÉRARD, 1972, GU *et al.*, 2011), and renewed by Romney and colleagues (ROMNEY *et al.*, 1996, 2000), who introduced a label-free version of such dimensional analysis that allows the creation of what they call the “shared cognitive structures” (ROMNEY; MOORE, 1998; cf. RILLIARD *et al.*, 2014 for an application on prosodic meaning). The eleven first dimensions of the correspondence analysis (which account for 70% of the total variance) were kept, based on an elbow criterion. The dispersion of the 34 types of expressions (17 attitudes on 2 sentences) on these eleven dimensions was taken as the perceived dispersion of the 34 types of expressions, in a space based on these 274 labels; Husson *et al.* (2011, p. 189) defend the idea that a clustering based on the main dimensions of a multidimensional analysis removes noisy data and gives more stable partitions. A hierarchical clustering was thus run on

the eleven first principle components of the correspondence analysis (using the HCPC procedure of the *FactoMineR* library; HUSSON *et al.*, 2011). This allows the creation of a dendrogram showing the relative similarity of the 34 types of expressions (cf. figure 1). At the main levels of this tree, one may observe three large clusters that regroup respectively the expressions of (i) ADMI, SEDU, SURP; (ii) ARRO, AUTH, CONT, DECL, IRRI, OBVI, POLI, SINC, and IRON on the B sentence; and (iii) DESC, DOUB, QUES, UNCE, WOEG, and IRON on the M sentence. Cluster (i) is mostly described (by the labels representing more than 5% of the occurrences of labels in this cluster) using the labels “*surpresa, alegria, admiração*” (surprise, joy, admiration): it is thus linked to behaviours carrying notions of novelty and positive valence. Cluster (ii) is described as “*afirmação, obviedade, certeza*” (affirmation, obviousness, certainty): it is thus linked to notions of assertion and assertive behaviours. Cluster (iii) is described as “*dúvida, incerteza*” (doubt, uncertainty): it is thus a cluster evoking dubitative expressions. The tri-partition of the 34 expressions recalls two main dimensions of meanings discussed respectively by Osgood *et al.* (1975) and Brandt (2008): a valence component, and the linguistic distinction between assertive and interrogative (or dubitative) speech acts.

In order to have a better understanding of the notions carried by these expressions, a finer level of cluster may be used. By selecting an appropriate level, according to a criterion of inertia gain (HUSSON *et al.*, 2011, p. 188), one may cut that tree so to select a set of clusters that maximizes their internal coherence and external difference. This procedure gives 12 distinct clusters grouping the 34 types of expressions, to be analysed in the next section.

FIGURE 1 – Dendrogram showing the relative distance between each 34 types of expressions (attitude, sentence) on the 11 first principal dimensions of the correspondence analysis, according to a hierarchical clustering procedure



Legend: The horizontal red line indicates the difference level selected to cut the tree so to obtain 12 clusters.

3 Analysis

3.1 Perceived dimensions in performances

The composition of the twelve clusters obtained from the free-labelling experiment is detailed in table 1. Half of the clusters (6) are composed of the two versions of the same expression, in each of the two sentences. These clusters (#2, #3, #5, #6, #8, #9) contain expressions that are distinguished from the others, whoever do perform them, and for the two situations that were used to elicit them. These expressions correspond to the situation of (in the same order than the clusters): IRRI, CONT, WOEG, UNCE, QUES, and DOUB.

TABLE 1 – List of the expressions (attitude, sentence) regrouped in each of the 21 clusters obtained in the free-labelling experiment

Cluster	Pairs of (Attitude, Sentence) composing the cluster				
#1	(ARRO, B) (DECL, M)	(ARRO, M) (POLI, B)	(AUTH, B) (POLI, M)	(AUTH, M) (SINC, B)	(DECL, B) (SINC, M)
#2	(IRRI, B)	(IRRI, M)			
#3	(CONT, B)	(CONT, M)			
#4	(OBVI, B)	(OBVI, M)	(IRON, B)		
#5	(WOEG, B)	(WOEG, M)			
#6	(UNCE, B)	(UNCE, M)			
#7	(DESC, B)	(DESC, M)	(IRON, M)		
#8	(QUES, B)	(QUES, M)			
#9	(DOUB, B)	(DOUB, M)			
#10	(SEDU, B)				
#11	(ADMI, M)	(SEDU, M)			
#12	(ADMI, B)	(SURP, B)	(SURP, M)		

Among the other clusters, one may find the following situations:

- Clusters with one expression recognized on both sentences, but also containing another expression. This is the case of clusters #4, #7 and #12, respectively composed of OBVI with IRON on the B sentence; DESC with IRON on the M sentence; and SURP with ADMI on the B sentence.
- Cluster 1, mixing many expressions together.
- Cluster 10, grouping only expressions of seduction on the “B” sentence.
- Cluster 11, grouping expressions of admiration and seduction on the M sentence.

To understand these clustering of expressions, one has to look at the list of labels used significantly more frequently inside each of these clusters than in the whole picture. These labels give a description of how the listeners have interpreted each cluster. Table 2 gives this list of labels describing the main expression perceived for the expressions composing each cluster. For the analysis, we'll focus on the more important labels: those set in bold in table 2 (with a criterion of composing more than 5% of the total of labels in that cluster)—of course, the complete list of labels participates in the characterisation of the expressions.

The five “homogeneous” clusters (those composed of only one attitude) are described as follows. Cluster #2, composed of expressions of IRRI, is coined as “*impaciência, raiva, irritação*” (impatience, anger, irritation): it is thus possible to conclude the expression is indeed well recognized by listeners, and that these expressions carry semantic traits of urgency (*impaciência*), negative valence (*raiva, irritação*), and high arousal (*raiva*). Cluster #3, composed of expressions of CONT, is described as “*nojo, desprezo, desgosto, indiferença*” (disgust, contempt, displeasure, indifference), and is also well recognized, and characterized by a negative valence, and a low arousal. Cluster #5 contains the expression coined as WOEG that corresponds to the Japanese *kyoshuku*, and does not have a direct translation in BP. It is described by subjects as “*vergonha, timidez, medo*” (shame, shyness, fear), and one can link these terms with Sadanobu’s definition (2004, p. 34; cf. supra), that contains the notion of shame and traits of submission. The long list of terms in this cluster (22) also shows the difficulty listeners have to coin a precise term on this situation—even if they do describe it rather accurately. The labels share traits of low arousal, negative valence and submissive behaviour.

TABLE 2 – List of the labels observed significantly more frequently inside the cluster than their global distribution (according to a v test, cf. HUSSON *et al.*, 2011)

Cluster	Labels (percentage of observation inside the cluster)
#1	afirmação (19,5%), certeza (13,2%), neutralidade (5,6%), confirmação (9,5%), constatação (3,6%), pedido (2,5%), segurança (2,0%), resposta (1,7%), normalidade (1,9%), atestação (0,9%), informação (1,4%), diretamente (0,8%), seriedade (1,3%), declaração (1,3%), assertividade (0,7%), convicção (0,8%), ordem (1,1%), inexpressivo (0,9%), solicitação (0,5%), narração (0,5%), arrogância (1,0%)
#2	impaciência (17,0%), raiva (11,4%), irritação (5,7%), agressividade (3,4%), nervosismo (2,3%), opressão (2,3%), ênfase (3,4%), exaltamento (1,7%), ordem (2,8%), correção (2,3%), desaprovação (2,3%), grosseria (1,1%), insistência (1,1%)
#3	nojo (11,9%), desprezo (12,5%), desgosto (6,3%), indiferença (5,7%), rejeição (2,3%), inveja (2,3%), arrogância (3,4%), repugnância (1,7%), incômodo (2,3%), desdém (4,0%), “tanto faz” (1,1%), insatisfação (1,1%), descaso (1,1%), desânimo (1,1%), raiva (4,0%), impaciente (1,1%)
#4	obviedade (42,4%), confirmação (9,8%), reafirmação (1,5%), sarcasmo (2,3%), explicação (1,5%), deboche (2,7%), simplicidade (0,8%), afirmação para pergunta idiota (0,8%)
#5	vergonha (10,2%), timidez (5,7%), medo (5,7%), constrangimento (3,4%), hesitação (4,5%), mentira (3,4%), tristeza (4,0%), insegurança (3,4%), lamentação (2,8%), incerteza (8,5%), decepção (2,8%), receio (1,7%), excitação (1,7%), incômodo (2,3%), desconforto (1,1%), confissão (1,1%), admitindo (1,1%), acanhamento (1,1%), descontentamento (1,7%), recordação (1,1%), inexistência (1,1%), cautela (1,1%)
#6	incerteza (33,0%), dúvida (36,4%), insegurança (4,0%), hesitação (2,8%), reprovação (1,7%), palpite (1,7%)
#7	ironia (24,6%), descrédito (5,3%), incredulidade (5,7%), deboche (4,9%), incredibilidade (4,2%), desdém (4,9%), sarcasmo (2,7%), negação (1,5%), decepção (1,9%), discordância (1,1%), descrença (1,9%)
#8	dúvida (41,5%), pergunta (14,2%), interrogação (9,1%), questionamento (4,5%), palpite (2,8%), desinteresse (1,7%), desconhecimento (1,7%), indagação (1,1%), apatia (2,3%)
#9	estranheza (10,2%), estranhamento (10,2%), dúvida (25,0%), descrença (5,7%), incredibilidade (6,3%), desconfiança (4,5%), incredulidade (5,1%), descrédito (4,0%), indignação (1,7%), interrogação (3,4%)

#10	malícia (6,8%), sedução (5,7%), felicidade (8,0%), desejo (4,5%), satisfação (5,7%), paixão (4,5%), gracejo (2,3%), contentamento (4,5%), alegria (6,8%), vontade (2,3%), sugestão (2,3%), gostar (2,3%)
#11	admiração (18,2%), encantamento (8,5%), contentamento (5,7%), alegria (8,5%), paixão (2,8%), apreciação (1,7%), felicidade (4,0%), lembrança (2,3%), prazer (1,7%), memória (1,7%), saudade (1,7%), elogio (1,7%), correção (2,3%), impressionado (1,1%), deleite (1,1%), fascínio (1,7%), ironia (8,0%), encanto (1,1%), entusiasmo (1,7%)
#12	surpresa (56,4%), alegria (8,7%), espanto (4,5%), descoberta (2,7%), felicidade (3,8%), empolgação (1,1%), entusiasmo (1,9%), satisfação (2,3%), maravilhado (1,1%), deslumbramento (0,8%), agrado (0,8%), fascínio (1,1%), encantamento (2,3%), exclamação (0,8%)

Note: Labels are listed in decreasing order of the test's importance; the percentage of observations inside clusters are reported.

Cluster #6, based on the expressions of UNCE, is labelled as “*incerteza, dúvida*” (uncertainty, doubt): this expression is clearly recognized, and labelled accurately by listeners. Cluster #8, based on the expressions of neutral QUES, is labelled as “*dúvida, pergunta, interrogação*” (doubt, question, interrogation): once again, this expression is recognized and labelled rather accurately, if the main label (doubt) is not the most typical of what a linguistic description would say. Cluster #9, based of expressions of DOUB, is labelled as “*estranheza, estranhamento, dúvida, descrença, incredibilidade*” (strangeness, estrangement, doubt, disbelief, incredibility): this show the quality of the recognition listeners made of these expressions. The three clusters #6, #8 and #9 share similar traits: these three expressions (QUES, UNCE, DOUB) share the same labels of “*dúvida*” (doubt), and traits related to doubt, like a submissive behaviour, and low activation levels. This correspond to interrogatives as described by Ohala (1994) when he proposes the linguistic interpretations of his Frequency Code, thus one may expect relatively higher levels of pitch for the expressions related to these clusters.

This interpretation differs from that of the cluster #7 (DESC and IRON on sentence M), described as “*ironia, descrédito, incredulidade*”. These terms carry a rejection of an assertion, as in cluster #9 (“*descrença*”), but without a component of doubt; on the contrary, an ironic tone is clearly noted, together with labels like “*sarcasmo, desdém,*

discordância” (sarcasm, disdain, disagreement). The interpretation of this cluster is more assertive (in the rejection of the assertion) than cluster #9.

Cluster #4 contains the OBVI expression, as well as the expression of IRON on sentence B. It is described as “*obviedade, confirmação*” (obviousness, confirmation): thus the listeners clearly interpret these expressions in a manner coherent with the situation of OBVI. The situation of IRON for sentence B consists in choosing between a banana and a sport car—with the speaker speaking “banana”, expressing obviously s/he prefers the reverse choice. The irony only appears from the context (the switch between the obviousness of the tone and the lexical content, regarding the context of choice), context which was not accessible to subject of the free labelling experiment.

Cluster #10, #11, and #12 are respectively described by the labels “*malícia, sedução, felicidade, satisfação, alegria*” (malice, seduction, happiness, satisfaction, joy), “*admiração, encantamento, contentamento, alegria*” (admiration, enchantment, contentment, joy), and “*surpresa, alegria*” (surprise, joy). These three clusters show similarities around a set of labels carrying positive valence (joy, happiness). Cluster #10 contains the expression SEDU in the context of B sentence, which is indeed described as a joyful play to seduce the interlocutor; the second situation expressing seduction was not recognized as such, but is mixed, in cluster #11, with ADMI on sentence M: both expressions share traits expressing the quality of an object of desire, but with a more passive pattern than SEDU on the B sentence. This reflects adequately the differences in both situations aiming at eliciting SEDU on the B and M sentences: for the B sentence, the speaker has to express her/his sexually related interest to the interlocutor; for the M sentence, the speaker expresses her/his feeling about the interlocutor. The second situation of admiration (on the B sentence) is mixed with the two expressions of SURP: they share a trait of surprise (and admiration on the B sentence was elicited with a kind of surprise—the fruit being shown suddenly), and a positive valence.

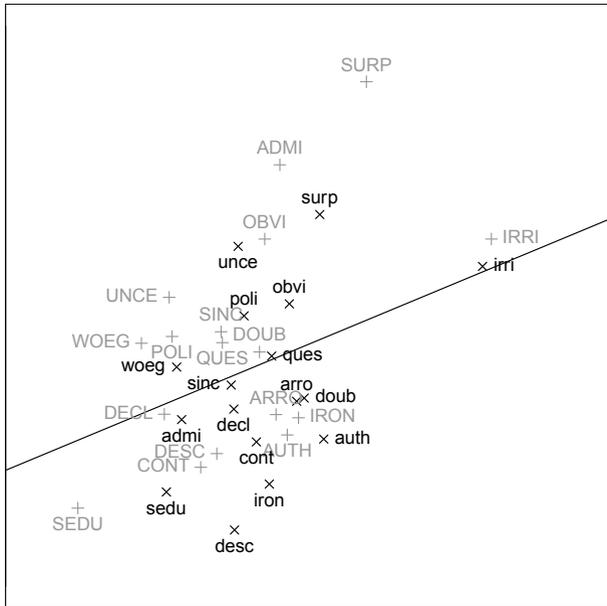
Finally, cluster #1 regroups the largest set of expressions: ARRO, AUTH, DECL, POLI and SINC. All these expressions share an assertive sentence mode, and are mostly described following this trait, as “*afirmação, certeza, neutralidade, confirmação*” (affirmation, certainty, neutrality, confirmation). The traits fit adequately to the neutral declarative situation, but they fail to describe adequately the specificities of the other expressions (ARRO, AUTH, POLI, SINC):

these four expressions thus lead to performances that are not adequately perceived as such (out of context) by the listeners. At least, it is not the prominent characteristic of the corresponding speech acts. Meanwhile, digging up a little in the tree (cf. figure 1), one may observe that the next level of clustering associated to this cluster separates ARRO and AUTH from DECL, POLI and SINC. There seems to be some information on the valence of the speech act and/or its imposition degree, but it is not something that characterizes the performances prominently.

3.2 Acoustical variations

In order to observe the acoustic variation across the prosodic performances of these 17 attitudes, several parameters have been extracted (cf. supra). Figure 2 shows the dispersion of attitudes along the parameters of intensity and F_0 . Both have been reported to play an important role in expressive voice. F_0 is the most prominent acoustic correlate of prosodic variations, being mostly related to the perception of pitch, while intensity is correlated to vocal effort (LIENARD; BARRAS, 2013). Changes in both parameters are constrained by the voice production mechanism, and an increase in vocal effort is generally linked to a rise in F_0 (TITZE; SUNDBERG, 1992). Changes in vocal effort have been reported to be the primary acoustic cues in the expression of vocal emotion, and are related to expressive arousal (BANSE; SCHERER, 1996; GOUDBEEK; SCHERER, 2010).

FIGURE 2 – Position of the mean values of each type of expression (see text for abbreviation) on the intensity $\times F_0$ plane (both expressed as z-score), as performed within each of the two sentences



Legend: B: uppercase grey fonts;

M: lowercase black font.

The regression line shows the relation between both parameters.

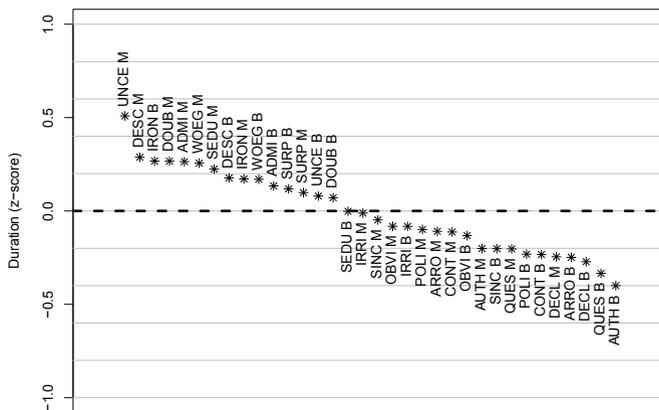
The dispersion of expressions observed on figure 2 illustrates these tendencies: expressions belonging to clusters described with terms linked to high activation (typically cluster #2, with expressions of IRRI) show the highest levels of intensity, and also relatively high F_0 —but this rise of F_0 follows the regression line between both parameters, and shall be the sub-product of a louder voice, and hence related to the Effort Code. The spread of expressions along a direction perpendicular to the regression line follows changes in F_0 that are not explained by change in vocal effort, and hence related to the Frequency Code. This line separates expressions with a higher F_0 , above the line, from expressions with lower F_0 , under the line—for comparable levels of intensity. Such a separation allows to regroup expressions of WOEG, UNCE, QUES,

POLI, SINC, marked by semantic features of submission or positive valence, and performed with a higher pitch, from expressions of DESC, CONT marked by assertive or negative features.

Differences between the two types of elicitations of the same expression are also observed. SEDU on the B sentence received the lowest values of intensity and a low F_0 . It thus departs clearly from other expressions, and this low voice may be related to a search of intimacy linked to the expression of seduction. On the contrary, SEDU on the M sentence still show a low pitch, but levels of intensity closer to those of the expressions of ADMI it is mixed with. Interestingly, this type of ADMI on the M sentence (performed with F_0 and intensity values close to those of declaration) also departs clearly from the sentence B type of ADMI, which shows very high F_0 , for a relatively medium intensity level: in that respect, it is acoustically comparable to the expressions of SUPR it is mixed with.

On the middle of the graph lie the many expressions that share conversational levels of intensity and F_0 . Typically, one observes the proximity of the expressions regrouped under cluster #1 that share similar values of mean pitch and vocal effort.

FIGURE 3 – Mean standardized syllabic duration (z-score) of each type of expression, on each sentence (see text for abbreviation)



Legend: The dashed horizontal line indicates the mean syllabic duration.

Duration constitutes a third set of prosodic indices: mean syllabic durations are displayed for each expression on figure 3. Two groups of expressions may be observed: (i) expressions with a tendency to lengthening, that regroups (in decreasing order of lengthening) clusters #6, #7, #5, #11, #9, and #12, plus IRON on the B sentence; and (ii) expressions with a tendency to shortening, that regroups (in decreasing order of shortening) clusters #1, #8, #3, #4 (minus IRON), #2, and #10. These two groups are mostly differentiated by traits of dominance and assertiveness (linked to shortening) vs. submissiveness and dubitative expressions (linked to lengthening)—the case of QUES (cluster #8) being a counter-example. One may also observe that neutral sentence (both DECL and QUES) are performed with fast speech rhythm, while most expressive behaviours do exhibit some lengthening (i.e. all expressions show lengthened syllabic durations regarding these two expressions but those of cluster #1). The two expressions of SEDU show different patterns of lengthening, with SEDU on the M sentence, grouped with ADMI (and described as a kind of admiration), having the longest—while SEDU on the B sentence have a faster, more assertive rhythm. Expressions of IRON on the B sentence are grouped with lengthened expressions and not with expressions of OBVI (cluster #4); such difference in timing may have been used by speakers as a cue for expressing exaggeration (a component of ironic meaning, according to BRYANT, 2011; GONZÁLEZ-FUENTE, 2015), but appears to be insufficient for the listeners being able to decode them as ironic, outside of the communication context.

4 Discussion and conclusions

The results of the free-labelling study have shown that the 17 expressions are first regrouped in three main groups, which are interpreted in terms of: (i) positive valence, (ii) assertive expressions, and (iii) dubitative expressions. These distinctions are basic components of meaning: valence is part of the three “dimensions of meaning” cross-culturally observed by Osgood et al. (1975); the distinction between assertion / dubitative expressions being a classical dimension of linguistic meaning, and one of the basic functions of prosody (the expression of sentence’s mode; BRANDT, 2008). It is worth noting that only one of the three “dimensions of meaning” reported by Osgood is part of these three main distinctions, and that the next ones (Osgood’s dimensions of

activation and dominance) are not part of this high-level description. The fact that Osgood's work was based on isolated (and written) words may explain why the typically interactional distinction between assertive and interrogative speech acts was not observed in his work. This distinction was also observed in works on other languages performed with the same methodology as the one presented here, typically in French, Japanese (cf. GUERRY *et al.*, 2015, 2016; other languages are under study): it even constitutes the main distinction found in these two languages, for which the valence component do not seem to be as prominent (a comparative study shall be pursued to raise conclusive remarks on this aspect). This distinction between assertion and interrogation is also reflected in the acoustic measures, and ranked along the regression line drawn from the position of vowels on the intensity x F_0 plane. This distinction, orthogonal to the line linked to vocal effort, can be interpreted as a change in F_0 explicitly made by the speaker to change the pitch of her/his production (and not as a by-product of a stronger vocal effort): this is in line with the interpretation proposed by the Frequency Code (OHALA, 1994).

At a finer level of distinctions, 12 clusters have been described. Among these 12 groups of expressions, half are based on one type (and one only) of attitude among the 17, and labelled with terms that correspond to the intended expressions. There are thus 6 expressions that are recognized without ambiguity, and not mixed with others: IRRI, CONT, WOEG, UNCE, QUES, and DOUB. Three more clusters are based on a well-recognized expression, but also contain another expression that was not labelled adequately (in relation to the targeted attitude by the listeners: OBVI (mixed with IRON on the B sentence), DESC (mixed with IRON on the M sentence) and SURP (mixed with ADMI on the B sentence).

ADMI is recognized in one case (on the M sentence), and mixed with one type of seduction, that involve prizing the interlocutor; the second type of seduction is recognized and labelled according to the speaker's intended attitude. The accuracy of the listeners' descriptions is interesting as they manage to grasp the difference in the scenario set up to elicit most of the 17 attitudes—with the notable exception of four expressions part of cluster #1. This cluster regroups a set of assertive expressions that have not been singled-out by the listeners: ARRO, AUTH, POLI, and SINC (cluster #1 being mostly labelled as a neutral assertion, one may conclude DECL is well recognized).

One may compare the recognition of the complete set of 16 expressions (let's remember for other language, the DESC situation was not recorded) in PB with the results obtained in French (GUERRY *et al.*, 2015) and Japanese (GUERRY *et al.*, 2016). French listeners (judging French expressions) have made the same grouping of IRON on the B sentence with OBVI that the one observed here (the expression of IRON on the M sentence being mixed with negative and dominant expressions by French); on the contrary, both expressions of IRON are well clustered in the Japanese dataset, and interpreted as a negative sarcasm. This notion of sarcasm can be compared with labels given to IRON on the M sentence by both Brazilian and French listeners. Another common behaviour is related to the comparison of SEDU with ADMI: both Brazilian and French listeners do mix these two expressions in the context of the M sentence only (associating them to an expression of admiration), while Japanese subjects do these grouping for all expressions of ADMI and SEDU; conversely, French and Brazilian subjects do note the sexually related expression of SEDU in the B sentence (French do mix this one with ADIM on the B sentence, with labels expressing desire and longing).

The situation corresponding to the Japanese *kyoshuku*, WOEG, is well distinguished by Japanese subjects, as for Brazilian, and described primarily as “申し訳ない” (I'm sorry) by Japanese and as “*vergonha*” (shame) by Brazilian. In both cases, this expression pertains to a more general dubitative cluster. Let's note that even for Japanese, who did conventionalize this expression in their language, the list of terms used to denote it is particularly long (25 in the case of Japanese, 22 for Brazilian). French subjects do not singularize the WOEG expression, but mixed it with UNCE, and do label it accordingly.

A singularity of the Brazilian data lies in the large set of clusters (12) compared to seven for French and eight for Japanese. This shows a greater accuracy in singularizing a set of unique expressions, compared to the two other languages. Finally, both French and Japanese do separate a large group of negative and/or imposing expressions (AUTH, IRRI, CONT, ARRO) from a neutral or positive group of assertions based on DECL, and mixed with POLI and SINC. This is not the case for Brazilians, who do mix these two groups, singularizing IRRI and CONT, but neither AUTH nor ARRO from DECL. Note the valence component is present (and important) in the Brazilian data, with a group of expressions with a positive valence. Once again, a compared analysis shall be pursued

for interpreting these observations and propose a comprehensive cross-cultural analysis. To that aim, a detailed analysis of the semantic features linked to each type of expressions shall be pursued. It will notably help a better understanding of the various prosodic expressions captured in this corpus, and to propose a set of comparable expressions across languages, as well as another set of varying expressions, worth to be studied in the framework of foreign language teaching.

Acknowledgements

The authors are deeply indebted to the many speakers and listeners from Rio de Janeiro Federal University, who participated in these recordings and perception experiments. Also, we acknowledge the French National Research Agency (ANR) PADE, and the Brazilian CNPq agency (bolsa PV 406177/2015-5).

References

- BANSE, R.; SCHERER, K. R. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, Washington, American Psychological Association, v. 70, n. 3, p. 614-636, 1996.
- BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer* [Computer program]. Version 6.0.07. Disponível em: <<http://www.praat.org/>>. Acesso em: 03 sep. 2016.
- BOROD, J. C.; PICK L. H.; HALL, S.; SLIWINSKI, M.; MADIGAN, N.; OBLER, L. K.; WELKOWITZ, J.; CANINO, E.; ERHAN, H. M.; GORAL, M.; MORRISON, C.; TABERT, M. Relationships among facial, prosodic, and lexical channels of emotional perceptual processing. *Cognition & Emotion*, Taylor & Francis Online, v. 14, n. 2, p. 193-211, 2000.
- BRANDT, P. A. Thinking and language. A view from cognitive semio-linguistics. In: INTERNATIONAL CONFERENCE ON SPEECH PROSODY, 4th, Campinas, Brazil. *Proceedings...* Campinas: Unicamp, 2008. p. 649-654.
- BRYANT, G. A. Verbal irony in the wild. *Pragmatics & Cognition*, John Benjamins, v. 19, n. 2, p. 291-309, 2011.

CHANG, N. C. T. Tones and intonation in the Chengtu dialect (Szechuan, China). *Phonetica*, Karger Publishing, v. 2, n. 1-2, p. 59-85, 1958.

DANEŠ, F. Involvement with language and in language. *Journal of Pragmatics*, Elsevier, v. 22, n. 3-4, p. 251-264, 1994.

DELATTRE, P. Comparing the prosodic features of English, German, Spanish and French. *International Review of Applied Linguistics in Language Teaching*, De Gruyter, v. 1, n. 1, p. 193-210, 1963.

FÓNAGY, I.; BÉRARD, E. «Il est huit heures»: contribution à l'analyse sémantique de la vive voix. *Phonetica*, Karger Publishing, v. 26, n. 3, p. 157-192, 1972.

FUJISAKI, H.; HIROSE, K. Analysis and perception of intonation expressing paralinguistic information in spoken Japanese. In: ESCA WORKSHOP ON PROSODY, 1993, Lund, Sweden. *Proceedings...* Lund, Sweden: University of Lund, 1993. p. 254-257.

GOLDMAN, J. P. EasyAlign: an automatic phonetic alignment tool under Praat. In: ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION (InterSpeech), 12th, 2011, Firenze, Italy. *Proceedings...* Firenze, Italy: International Speech Communication Association, 2011. p. 3233-3236.

GONZÁLEZ-FUENTE, S.; ESCANDELL-VIDAL, V.; PRIETO, P. Gestural codas pave the way to the understanding of verbal irony. *Journal of Pragmatics*, Elsevier, v. 90, p. 26-47, 2015.

GOUDBEEK, M.; SCHERER, K. R. Beyond arousal: Valence and potency/control cues in the vocal expression of emotion. *The Journal of the Acoustical Society of America*, Acoustical Society of America, v. 128, n. 3, p. 1322-1336, 2010.

GREENBERG, Y.; SHIBUYA, N.; TSUZAKI, M.; KATO, H.; SAGISAKA, Y. Analysis on paralinguistic control in perceptual impression space using multiple dimensional scaling. *Speech Communication*, Elsevier, v. 51, p. 585-592, 2009.

GU, W.; ZHANG, T.; FUJISAKI, H. Prosodic Analysis and Perception of Mandarin Utterances Conveying Attitudes. In: ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION (InterSpeech), 12th, 2011, Firenze, Italy. *Proceedings...* Firenze, Italy: Speech Communication Association, 2011. p. 1069-1072.

GUERRY, M.; SHOCHI, T.; RILLIARD, A.; ERICKSON, D. Perception of prosodic social affects in French: a free-labeling study. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES (ICPhS), 18th, 2015, Glasgow, UK. *Proceedings...* Glasgow, UK: The Scottish Consortium for ICPhS, 2015. 4p.

GUERRY, M.; RILLIARD, A.; ERICKSON, D.; SHOCHI, T. Perception of prosodic social affects in Japanese: A free-labeling study. In: SPEECH PROSODY, 8th, 2016, Boston, USA. *Proceedings...* Boston, USA: University of Boston, 2016. p. 811-815.

GUSSENHOVEN, C. *The phonology of tone and intonation*. Cambridge: Cambridge University Press, 2004.

HANSON, H. M. Glottal characteristics of female speakers: Acoustic correlates. *The Journal of the Acoustical Society of America*, Acoustical Society of America, v. 101, n. 1, p. 466-481, 1997.

HUSSON, F.; LÊ, S.; PAGES, J. *Exploratory Multivariate Analysis by Example Using R*. London: Chapman and Hall, 2011.

LIENARD, J. S.; BARRAS, C. Fine-grain voice strength estimation from vowel spectral cues. In: ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION (InterSpeech), 14th, 2013, Lyon, France. *Proceedings...* Lyon, France: Université de Lyon, 2013. p. 128-132.

MARTINS-BALTAR, M. *De l'énoncé à l'énonciation: une approche des fonctions intonatives*. Paris: Didier, 1977.

OHALA, J.J. The frequency codes underlies the sound symbolic use of voice pitch. In: HINTON, L.; NICHOLS, J.; OHALA, J. J. (Ed.). *Sound symbolism*. Cambridge: Cambridge University Press, 1994. p. 325-347.

OSGOOD, C. E.; SUCI, G. J.; TANNENBAUM, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.

OSGOOD, C. E.; MAY, W. H.; MIRON, M. S. *Cross-cultural universals of affective meaning*. Urbana: University of Illinois Press, 1975.

R CORE TEAM. *R: A language and environment for statistical computing*. [Computer Program]. Vienna, Austria: R Foundation for Statistical Computing. Disponível em: <<https://www.R-project.org/>>. 2016.

RILLIARD, A.; ERICKSON, D.; SHOCHI, T.; MORAES, J. A. Social face to face communication – American English attitudinal prosody. In: ANNUAL CONFERENCE OF THE INTERNATIONAL SPEECH COMMUNICATION ASSOCIATION (InterSpeech), 14th, 2013, Lyon, France. *Proceedings...* Lyon, France: Université de Lyon, 2013. p. 1648-1652.

RILLIARD, A.; ERICKSON, D.; DE MORAES, J. A.; SHOCHI, T. Cross-cultural perception of some Japanese politeness and impoliteness expressions. In: BAIDER, F.; CISLARIU, G. (Ed.). *Linguistic approaches to emotions in context*. Amsterdam: John Benjamins Publishing Company, 2014. p. 251-276.

ROMNEY, A. K.; BOYD, J. P.; MOORE, C. C.; BATCHELDER, W. H.; BRAZILL, T. J. Culture as shared cognitive representations. *Proceedings of the National Academy of Sciences*, National Academy of Sciences of the United States of America (PNAS), v. 93, n. 10, p. 4699-4705, 1996.

ROMNEY, A. K.; MOORE, C. C. Toward a theory of culture as shared cognitive structures. *Ethos*, Wiley Online Library, v. 26, n. 3, p. 314-337, 1998.

ROMNEY, A. K.; MOORE, C. C.; BATCHELDER, W. H.; HSIA, T.-L. Statistical methods for characterizing similarities and differences between semantic structures. *Proceedings of the National Academy of Sciences*, National Academy of Sciences of the United States of America (PNAS), v. 97, n. 1, p. 518-523, 2000.

SADANOBU, T. A natural history of Japanese pressed voice. *Journal of the Phonetic Society of Japan*, Phonetic Society of Japan, v. 8, n. 1, p. 29-44, 2004.

SCHERER, K. R. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, Royal Society Publishing, v. 364, n. 1535, p. 3459-3474, 2009.

SPENCER-OATEY, H. Reconsidering power and distance. *Journal of Pragmatics*, Elsevier, v. 26, p. 1-24, 1996.

TITZE, I. R.; SUNDBERG, J. Vocal intensity in speakers and singers. *The Journal of the Acoustical Society of America*, Acoustical Society of America, v. 91, n. 5, p. 2936-2946, 1992.

TRAUNMÜLLER, H.; ERIKSSON, A. Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America*, Acoustical Society of America, v. 107, n. 6, p. 3438-3451, 2000.

ULDALL, E. Attitudinal meanings conveyed by intonation contours. *Language and Speech*, Sage Publications, v. 3, n. 4, p. 223-234, 1960.

WIDEN, S. C.; RUSSELL, J. A. A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, American Psychological Association, v. 39, p. 114-128, 2003.

WIERZBICKA, A. Different cultures, different languages, different speech acts: Polish vs. English. *Journal of pragmatics*, Elsevier, v. 9, n. 2-3, p. 145-178, 1985.

WU, Y. C.; COULSON, S. How iconic gestures enhance communication: An ERP study. *Brain and Language*, Elsevier, v. 101, n. 3, p. 234-245, 2007.

Annex

We give hereafter the details of the prototypical situations that have been used to describe the 17 prototypical situations and dialogues corresponding to the targeted attitudes (in each case, F1 is the recorded speaker, F2 the interlocutor):

Declarativo neutro	Neutral declarative sentence (DECL)
O falante F1 dá uma informação ao falante F2 (“Maria dançava” / “Uma banana”), de forma neutra, isto é, sem nenhuma atitude ou perspectiva pessoal. F1 & F2 são colegas, de mesma idade, e estão num local informal, p. ex. num bar.	F1 & F2 are colleagues, same age; F1 gives information without any personal perspective; the scene is at a coffee shop.
F2: O que ela estava fazendo quando você chegou? F1: Maria dançava.	F2: What was Mary doing when you arrived? F1: Mary was dancing.
Pergunta neutra	Neutral question (QUES)
Sem nenhuma expectativa particular, esperando uma resposta simples. Os Falantes F1 & F2 são amigos, mesma idade. Local: em um bar	F1 & F2 are colleagues, same age. F1 asks for information, without any personal perspective, awaiting a simple answer. The scene is at a coffee shop.
F2: Encontrei Maria ontem a noite. F1: Maria dançava?	F2: I saw Mary last night. F1: Mary was dancing?
Admiração	Admiration (ADMI)
F1 & F2 adoram comida japonesa, e eles falam de um prato delicioso que comeram em um restaurante japonês na semana passada. Ambos são aproximadamente da mesma idade, e se conhecem bem. No restaurante da universidade.	F1 & F2 are almost the same age and know each other well. Both love French cuisine, and talk about the very delicious food they ate yesterday at a famous French restaurant. The scene is at a coffee shop.
F2: E a comida japonesa que comemos na semana passada? F1: (Hummmmm...) Era maravilhosa!	F2: Do you remember the French food we ate yesterday? F1: Oh, yes, of course! It was so fantastic!

Arrogância	Arrogance (ARRO)
<p>O falante F1 está dando uma festa, e o falante B não foi convidado, mas A percebeu que ele /ela está na festa. Ambos, F1 & F2 estudam da mesma faculdade mas F1 é mais velho que o falante F2, seu pai é o reitor da Universidade e você é um tanto metido a snob. Os dois falantes se conhecem, mas não se gostam muito. Local: Na festa</p>	<p>Both F1 & F2 are from the same university, but F1 is older and F1's father is head of the university and F1 is a bit of a snob. Both know each other, but are not friends. F1 organized a social party, and F2 was not invited to the party, but F1 is aware of his/her presence during the party. The scene is a party room, and F1 says to F2 that only his friends are invited.</p>
<p>F1: O que que você está fazendo aqui? F2: É você que está dando essa festa? F1: (não), Eu só convido gente de alto nível para minhas festas.</p>	<p>F1: Why are you here? F2: You organized this party? F1: I only invite my friends to my parties.</p>
Autoridade	Authority (AUTH)
<p>O falante F2 está na frente do F1, requerendo permissão para entrar no país. F1 deve impor sua autoridade. O Falante F1 é um agente da polícia federal, o falante F2 é um turista. A cena se passa no controle de imigração do Aeroporto.</p>	<p>F1 is a custom agent; F2 is a traveller. F2 is in front of F1, requesting permission to enter the country; F1 needs to impose his authority; the scene is at a custom counter at the airport.</p>
<p>F1: Qual o motivo de sua visita? F2: Eu sou um estudante, fazendo intercâmbio F1: Por quanto tempo você vai ficar no país? F2: Três semanas</p>	<p>F1: What is your purpose for visiting? F2: I am an exchange student F1: How long do you plan to stay? F2: 3 weeks</p>
Desprezo	Contempt (CONT)
<p>O falante F1 está dando uma festa em casa e não convidou o falante F2, mas F1 descobre que ele está na festa. F1 e o falante F2 são colegas de faculdade, mas F1 é de um ano mais adiantado. Eles se conhecem, mas não são amigos. Na verdade F1 detesta F2. Durante a festa.</p>	<p>Both F1 & F2 are from the same university, but F1 is older; both know each other, but are not friends. In fact, F1 really hates F2. F1 organized a social party, and speaker F2 was not invited, but F1 is aware of his/her presence. The scene is at a party room</p>

<p>F1: Você não foi convidado, foi? (olhando para ele com ar superior) F2: Não, mas eu queria muito vir. F1: É melhor você ir embora!</p>	<p>F1: You are not invited, are you? (thinking about how you dislike F2 so much) F2: No, but I really wanted to come. F1: You should leave!</p>
Evidência	Obviousness (OBVI)
<p>Todo mundo sabe que Paulo não gosta de filmes do Woody Allen. Mas F2, um amigo de F1, ainda assim lhe pergunta se Paulo gostaria de ver um dos filmes do Woody Allen. Os dois falantes são colegas, mesma idade. Local: em um bar</p>	<p>F1 & F2 are colleagues, same age; everyone knows F2 doesn't like French movies, but F1 asks F2 if he likes French movies or not; the scene is at a coffee shop.</p>
<p>F2: Você acha que Paulo gostaria de ver um filme do Woody Allen hoje? F1: Acho que não!...</p>	<p>F2: Do you think Paul likes French movies? F1: Of course he doesn't like them.</p>
Incerteza	Uncertainty (UNCE):
<p>O falante F1 acha que viu Paulo, um amigo comum (dele e de seu interlocutor F2), no Maracanã na véspera, mas ele não está 100% seguro de que fosse realmente Paulo. Os dois falantes são colegas, mesma idade. Local: em um bar</p>	<p>F1 & F2 are colleagues, same age. F1 saw Paul at the baseball game yesterday, but is not 100% sure if it was really Paul; the scene is at a coffee shop.</p>
<p>F2: Oi, eu fui ao Maracanã ontem, o jogo foi ótimo. Você sabe se o Paulo também foi ao jogo? F1: Acho que o Paulo estava lá também...</p>	<p>F2: Hi, I went to see the baseball game yesterday, and it was really fantastic! Do you know if Paul went to see the game too? F1: Paul was probably there too.</p>
Ironia	Irony (IRON)
<p>O falante F1 está indo para São Paulo para assistir a um jogo do seu time e seu amigo Pedro F2, que mora em São Paulo telefona para ele. Infelizmente, o tempo em São Paulo está chuvoso. Os dois falantes são colegas, mesma idade. Local: no aeroporto do Galeão</p>	<p>F1 & F2 are friends, same age; F1 is going to Boston to see an important baseball game, and F2, who is living in Boston calls F1. Unfortunately, the weather in Boston is rainy and F1 says it's wonderful; the scene is at an airport.</p>
<p>F1: Oi, Pedro, como está o tempo hoje? F2: Hoje está chovendo... F1: Ah!, que ótimo!</p>	<p>F1: Hey Paul, what's the weather today? F2: Oh it's rainy today, unfortunately... F1: Oh! That's great!</p>

Irritação	Irritation (IRRI)
<p>O falante F1 está sentado ao lado do falante F2. O falante F2 acende um cigarro, e F1 fica muito irritado com isso. F1 quer que F2 pare de fumar, e expressa sua irritação em relação a ele. Os dois são aproximadamente da mesma idade, e se conhecem bem. No corredor da faculdade</p>	<p>F1 & F2 are almost the same age and know each other. F1 is sitting next to F2. Suddenly, F2 starts to smoke, and F1 is very angry; s/ he wants him/her to stop, expressing his irritation toward speaker F2. The scene is a public place.</p>
<p>F1: Por favor, apaga a cigarro. F2: Ok, Ok, eu sei.... F1: Por favor, apaga a cigarro!</p>	<p>F1: Excuse me, but don't smoke please. F2: ok I know I know... F1: Don't smoke, please!</p>
Pergunta com estranheza	Doubt (DOUB)
<p>O falante F1 sabe que seu amigo Paulo não sabe cozinhar, mas Paulo (Falante F2) conta ter feito um jantar ontem; você tem lá suas dúvidas se isso é mesmo verdade. Mas diferentemente da afirmação com descrédito (a anterior), o falante F1 estranha a afirmação feita por F2, e pergunta se é isso mesmo que ouviu, como se quisesse confirmar o que foi dito, ao mesmo tempo manifestando, por seu tom de voz, sua estranheza em relação a informação dada por F2. F1 & F2 são amigos, mesma idade. Num bar.</p>	<p>F1 & F2 are colleagues, same age. F1 knows that his colleague F2 didn't go to the baseball game yesterday, but F2 pretends he went to the game, and F1 doesn't believe it. The scene is at a coffee shop.</p>
<p>F2: Ontem eu fiz uma lasanha, ficou ótima F1: Você fez uma lasanha??? (tem certeza? Você não sabe fritar um ovo!)</p>	<p>F2: Hi, did you see the baseball game yesterday? It was really fantastic! F1: Hey Paul, did you really go to see the game yesterday?</p>

Pisando em ovos	Walking on eggs (WOEG)
<p>F2, o chefe de F1, quer que F1 se encarregue de uma tarefa muito trabalhosa, que vai exigir muitas horas fora do expediente e você sabe que não vai dar para fazer o que ele pede (sua mãe está no hospital e você precisa cuidar dela, mas você não quer dizer isso a seu chefe). Então você tenta recusar seu pedido tentando ao mesmo tempo não deixá-lo irritado com você por que está recusando a tarefa. O falante F2 é o chefe de F1, é mais velho que F1. No escritório de F2.</p>	<p>F2 is chief of the section which F1 belongs to; F2 is older than F1. The chief (F2) wants F1 to do a task which is a lot of work, and it seems to A it is impossible to do this, so F1 tries to reject this request by trying to make sure her/his boss (F2) doesn't get angry for refusing. The scene is at F2's office.</p>
<p>F2: Estou indicando você para se ocupar do próximo projeto. F1: Obrigado pela confiança, mas acho que não vou poder assumir essa responsabilidade. F2: Acho que pode, sim, você está recusando? F1: Desculpa, mas acho que não vou poder...</p>	<p>F2: I am recommending you to be in charge of our next big project. F1: Thank you for your confidence, but I'm afraid I can't take on such a big responsibility. F2: I think you can do it. Are you declining? F1: I'm sorry, but it seems difficult for me...</p>
Polidez	Politeness (POLI)
<p>F1 está sentado ao lado de F2. Ambos começam a conversar sobre banalidades, de forma muito educada. Ambos são aproximadamente da mesma idade. Eles não se conhecem muito bem, mas trabalham na mesma empresa. Numa festa da empresa.</p>	<p>F1 & F2 are almost the same age and don't know each other well, but work together professionally. F1 is sitting next to F1; both start social talk. The scene is at a formal party.</p>
<p>F1: Muito prazer F2: O prazer é meu... F1: Com é mesmo seu nome? F2: Meu nome é, e o seu?</p>	<p>F1: Nice to meet you. F2: I'm glad to meet you too. F1: Excuse me, would you remind me your name? F2: My name is, and you?</p>

Sedução	Seduction (SEDU)
<p>O falante F1 está apaixonado pelo falante F2. O Falante F1 elogia o falante F2 de forma provocativa, sensual, tentando atrair a atenção de F2. F1 & F2 tem uma relação íntima. Numa boate</p>	<p>F1 loves F2 and they have an intimate relationship. F1 gives a compliment to F2 in a sexually provocative way. The scene is at a clubhouse.</p>
<p>F1: Como você está sexy! F2: Obrigado. F1: Vamos dançar? F2: Claro, vamos.</p>	<p>F1: You look sexy today. F2: Thank you. F1: Would you dance with me? F2: Of course.</p>
Sinceridade	Sincerity (SINC)
<p>F2, o chefe de F1, quer que F1 se ocupe de um grande projeto. F1 está muito contente com a indicação, e expressa sua entusiasmo e sua vontade de cumprir bem essa tarefa. O Falante F2 é o chefe da seção onde F1 trabalha, e é mais velho que F1. Na sala de F2.</p>	<p>F2 is chief of the section which F1 belongs to; F2 is older than F1. The chief (F2) wants F1 to take on a big project; F1 is pleased to be asked to do this, and expresses his enthusiasm, honesty and sincerity for this task. The scene is at F2's office.</p>
<p>F2: Indiquei você para tomar conta do próximo projeto. F1: Muito obrigado pela confiança! Vou me esforçar ao máximo para não decepcioná-lo.</p>	<p>F2: I recommend you to be in charge of carrying out our next big project. F1: Thank you for your confidence. I will do my best. (appreciative)</p>
Surpresa	Surprise (SURP)
<p>O falante F1 não sabia que Paulo cantava tão bem. Uma dia, um amigo (Falante F2) faz você ouvir Paulo cantando. F1 & F2 são amigos, mesma idade. Na casa de F1.</p>	<p>F1 & F2 are friends, same age. F1 didn't know that F2 can sing well. One day, F2 makes F1 listen to his beautiful voice. The scene is at friend's home.</p>
<p>F2: Você já ouviu Paulo cantando? F1: Nunca ouvi. F2: A voz dele é incrível. Quer ouvir? F1: Claro. ***toca música*** F1: Nossa, é ele mesmo?</p>	<p>F2: Hey did you ever listen to Paul's song? F1: Not yet. F2: It's actually amazing. Do you wanna listen to his song? F1: Yes, sure. ***listen*** F1: Wow! Really?</p>

Descrédito	Discredit (DESC)
<p>O falante F1 sabe que seu amigo Paulo não sabe cozinhar, mas Paulo (Falante F2) insiste ter feito um jantar ontem; você não acredita. Diferentemente da ironia, aqui não se trata de uma resposta irônica, mas a repetição de uma afirmação anterior. O falante F1 repete o que ele acaba de ouvir, expressando sua falta de convicção em relação a informação dada por F2. Ele põe em causa ou mesmo duvida do que acaba de ouvir, mostrando por seu tom de voz que não acredita no que foi dito. F1 & F2 são amigos, mesma idade. Num bar.</p>	<p>F1 knows her/his friend Paulo (F2) cannot cook, but F2 insist in preparing the lunch; F1 disbelieve. Unlike for Irony, the answer here is not an ironic assertion, but the repetition of a preceding assertion. F1 repeats what s/he just heard, expressing her/his lack of conviction about the information given by F2. F1 rejects the plausibility of what s/he just heard, showing this disbelief in her/his tone of voice. Both F1 & F2 are friends of the same age. The scene is at a coffee shop.</p>
<p>F2: Ontem eu fiz uma lasanha, ficou ótima F1: Você fez uma lasanha (conta outra, fez nada!..)</p>	<p>F2: Yesterday, I cooked lasagne; it was great! F1: You cook lasagne (pull my leg)</p>

Perception of emotional prosody: investigating the relation between the discrete and dimensional approaches to emotions

Percepção da prosódia emocional: investigando a relação entre as abordagens discreta e dimensional das emoções

Wellington da Silva

Universidade Estadual de Campinas, Campinas, São Paulo / Brasil

wellingtondasilva@rocketmail.com

Plinio Almeida Barbosa

Universidade Estadual de Campinas, Campinas, São Paulo / Brasil

pabarbosa.unicampbr@gmail.com

Abstract: Emotional phenomena can be described according to various psychological approaches, the most adopted being the discrete (basic) and the dimensional ones. This study aimed at investigating the relation between the perception of some basic emotions and emotional dimensions in speech, as well as determining which acoustic cues are related to their perception. We conducted two perception experiments with utterances selected from a foreign language (Swedish) of which the listeners had no knowledge. In the first one, Brazilian subjects rated on 5-point scales the expressivity of four basic emotions: joy, anger, sadness, and calmness. In the second, a distinct group of Brazilian subjects rated the expressivity of five emotional dimensions: activation, fairness, valence, motivation, and involvement. The perception of the basic emotions and of the emotional dimensions was then compared by means of the Spearman's correlation coefficient. The five emotional dimensions were significantly correlated to some extent with the basic emotions, and these correlations were, in general, consistent with the literature and with the hypotheses that guided this study. We also performed an acoustic analysis,

in which twelve acoustic parameters were automatically computed for the utterances evaluated by the listeners. The parameters which correlated better with the listeners' judgments were fundamental frequency (median, interquartile semi-amplitude, 99.5% quantile), spectral tilt (mean and standard deviation), and LTAS slope. We concluded that it is possible to describe the perception of basic emotions in speech as a combination of emotional dimensions and that emotional dimensions may be better for describing the expression of emotions in speech.

Keywords: emotional prosody; basic emotions; emotional dimensions; perception test.

Resumo: Os fenômenos emocionais podem ser descritos de acordo com várias abordagens psicológicas, sendo a discreta (básica) e a dimensional as mais adotadas. Este estudo teve como objetivo investigar a relação entre a percepção de algumas emoções básicas e dimensões emocionais na fala, bem como determinar quais pistas acústicas estão relacionadas com sua percepção. Conduziram-se dois experimentos de percepção com enunciados selecionados de uma língua estrangeira (sueco) da qual os ouvintes não possuíam nenhum conhecimento. No primeiro, sujeitos brasileiros julgaram em escalas de cinco pontos a expressividade de quatro emoções básicas: alegria, raiva, tristeza e calma. No segundo, um grupo distinto de sujeitos brasileiros avaliou a expressividade de cinco dimensões emocionais: ativação, justiça, valência, motivação e envolvimento. A percepção das emoções básicas e das dimensões emocionais foi então comparada por meio do coeficiente de correlação de Spearman. As cinco dimensões emocionais correlacionaram-se significativamente em algum grau com as emoções básicas e essas correlações foram, no geral, consistentes com a literatura e com as hipóteses que nortearam este estudo. Realizou-se também uma análise acústica, na qual doze parâmetros acústicos foram computados automaticamente para os enunciados avaliados pelos ouvintes. Os parâmetros que melhor se correlacionaram com os julgamentos dos ouvintes foram: frequência fundamental (mediana, semi-amplitude entre quantis, quantil 99,5%), inclinação espectral (média e desvio padrão) e inclinação do LTAS. Concluiu-se que é possível descrever a percepção das emoções básicas na fala como uma combinação de dimensões emocionais e que as dimensões emocionais podem ser melhores para descrever a expressão de emoções na fala.

Palavras-chave: prosódia emocional; emoções básicas; dimensões emocionais; teste de percepção.

Recebido em: 17 de agosto de 2016.

Aprovado em: 16 de novembro de 2016.

1 Introduction

Emotions are complex phenomena which have intrigued researchers from different fields for decades. Many theories have been developed to explain what an emotion is (CORNELIUS, 2000). Currently, most scholars assume a componential view, for which emotions are (brief) processes made up of several components. This position differs, for example, from the one shared by the common sense, which sees emotion only as a feeling, that is, as the subjective experience of the state of emotional arousal caused by an event (SCHERER, 2000). The following components have been postulated (SCHERER, 2000): feeling, neurophysiological response patterns, motor expressions, action readiness, and appraisal. Mesquita & Frijda (1992) include in this list the antecedent events, event coding and regulation.

According to this perspective, an emotional process is triggered by the appraisal of an event in the environment (the antecedent event) as being relevant for the goals and needs of the individual. The appraisal of the event is automatic and unconscious, *i.e.*, it is not dependent on the deliberate action of the individual. However, it is mediated by the event coding, which is the meaning that the antecedent event has to the organism. Such process depends on the knowledge shared by the culture, the history of life, the personality and the personal beliefs of the person. Therefore, the emotion which the person will experience does not depend on the nature of the event itself, but rather on how it is coded by the individual (MESQUITA; FRIJDA, 1992). This is why the same event can trigger different emotions in different cultures and even between different people. Emotions are also subject to the regulation of the individual, who can try to disguise an expression (as is the case with shame) or exaggerate it (when someone tries to fake joy, for example). This process often depends on cultural and social norms, but it does

not inhibit the emotional process (for example, even disguising the expression, the person is still experiencing the emotion).

Researchers have also sought to find the most suitable way of describing the emotional processes and several methods have been developed for this purpose (COWIE; CORNELIUS, 2003). In the present study, we compare the two most influential approaches: the basic emotions approach and the dimensional. One of the big debates in emotion research is whether variations in the emotional processes should be regarded as gradual differences on a set of underlying dimensions or as discrete differences among emotion categories (FONTAINE, 2009).

According to the *basic emotions approach*, there exists a small set of universal emotions (known as basic or discrete), which are qualitatively different from each other and characterized by specific patterns of cognitive appraisal, expression, and physiological changes (EKMAN, 1992). This approach has its roots in the evolutionary theory of emotions (DARWIN, 2009[1872]), which understands that emotions are evolved phenomena which have been selected in the course of the evolution of the species because they help us to cope with relevant events in the environment, such as telling the individual when to attack, defend, flee, reject food, etc. As a result, one expects to find the same patterns of expression (facial and vocal) for a given emotion in all human cultures (CORNELIUS, 2000; MATSUMOTO; EKMAN, 2009). The basic emotions are described by means of labels provided by the natural languages, such as joy, sadness, anger, etc. Each of these labels actually denotes a family of related emotions. For example, the anger family is composed of emotions described by the labels angry, annoyed, irritated, furious, etc. (EKMAN, 1992; MATSUMOTO; EKMAN, 2009). However, there is no absolute agreement between researchers as to which emotions are “basic”. The most known proposal is Paul Ekman’s (EKMAN, 1992), which recognizes six basic emotions: joy, sadness, fear, disgust, anger, and surprise.

The *dimensional approach* understands that emotions are best represented by positions within continuous scales specified by underlying *emotional dimensions*. This approach, thus, emphasizes the gradual nature of the emotional phenomenon, which can occur with varying levels of intensity, and not only with two discrete poles of minimum and maximum intensity (FONTAINE, 2009; SCHLOSBERG, 1954). The most known emotional dimensions are *activation* and *valence*. Activation

(also labeled as arousal) refers to the degree of arousal of the individual (increase in behavior or physiological activity) and varies from calm to agitated (FOWLES, 2009; SCHLOSBERG, 1954). Valence (also known as pleasantness) corresponds to the subjective feeling of the degree of pleasantness caused by the antecedent event and emotions are commonly distinguished within this dimension as either positive (pleasant) or negative (unpleasant). Examples of positive emotions are joy and pride, and examples of negative emotions are anger, fear, shame (BROSCH; MOORS, 2009). It is possible to integrate the basic emotions approach with the dimensional by considering that each basic emotion can be represented by a set of underlying dimensions (MAUSS; ROBINSON, 2009; SMITH; ELLSWORTH, 1985). For example, the emotion labeled by the discrete approach as “anger” has negative valence and a degree of activation that varies from medium to high. The emotion “joy”, in turn, has positive valence, whereas “sadness” has negative valence and a low degree of activation.

In addition to valence and activation, many more emotional dimensions have been proposed in the literature (see, for example, FRIJDA *et al.*, 1995; SMITH; ELLSWORTH, 1985). Barbosa (2009) used the dimension of *involvement*, which is related to the degree of involvement of the individual with the event and is analogous to the opposition between attention – rejection used by Schlosberg (1941). A dimension that is also used is *dominance*, which refers to the capability of the individual to handle the situation, that is, whether he/she is in the control of the event or controlled by it (KEHREIN, 2002). Dominance has been regarded as an important dimension, as in some cases it is the only dimension that distinguishes pairs of emotions, such as anger from fear (SMITH; ELLSWORTH, 1985). However, some studies have found that dominance may not be well inferred from speech. Barbosa (2009) obtained very low inter-rater agreement for this dimension. Amir *et al.* (2010) obtained very low rates of automatic recognition (by means of combination of acoustic parameters) for the dimensions of dominance and valence, concluding, thus, that the acoustic parameters were better at predicting the dimension of activation than these other two dimensions. A similar result was found by Lugger & Yang (2007). Therefore, we decided here to replace dominance by other dimensions, in order to assess whether they can also be inferred from speech.

Following Frijda *et al.* (1995), we investigate in this study the perception of the dimensions of *fairness* and *motivation*. Fairness is related to the appraisal of the eliciting event by the individual, *i.e.*, whether the individual considered what happened fair or unfair. Motivation is a dimension related to action readiness, *i.e.*, whether the eliciting event enhanced or diminished the individual's disposition to act on the event. It has been suggested that the appraisal of an event as unfair can trigger and increase the intensity of various emotions, especially anger (ELLSWORTH; SCHERER, 2003, p. 581). In addition, the emotion the individual is expressing may also signal his/her behavioral intentions, such as approaching, avoidance, or touching, as well as his/her disposition to change his/her behavior (FRIJDA *et al.*, 1995).

Emotions can be expressed through different modalities, such as facial, body posture and gesture, and vocal (SCHERER; ELLGRING, 2007). This work addresses the expression of emotions in speech. The physiological responses triggered by emotions cause variations in respiration, phonation, and articulation, which are processes directly related to speech (SCHERER, 1986). In addition, the cognition of the speaker can also be affected by the emotional episodes, and this can have an impact on the temporal characteristics of speech (JOHNSTONE; SCHERER, 2000). Research on speech and emotion has focused on investigating the emotion-related changes in speech and whether listeners are able to recognize the intended emotion based only on speech samples (SCHERER, 2003). A common approach in this area is to measure acoustic properties of speech, such as amplitude, duration, and fundamental frequency, which we refer to in this work as "acoustic parameters". It is assumed that different emotions are characterized by specific patterns of changes in the acoustic parameters (see JOHNSTONE; SCHERER, 2000; and SCHERER, 2003, for a review of empirical findings on acoustic patterns for some basic emotions). The most studied parameters in the literature have been those related to vocal fold vibration (fundamental frequency), time (speech rate, duration of utterances and pauses) and intensity (perceived amount of energy in the speech signal). Because *prosody*, a branch of linguistics, is the area which studies these properties (BARBOSA, 2012), the term "emotional prosody" is often used to refer to utterances which convey emotional information.

With regard to the perception of emotional prosody, many studies which followed the basic emotion perspective have shown that listeners

can satisfactorily recognize a number of basic emotions through speech, even in intercultural contexts (*e.g.* BANSE; SCHERER, 1996; PITTAM; SCHERER, 1993; SCHERER; BANSE; WALLBOTT, 2001). Recent studies have demonstrated that some emotional dimensions can also be inferred from speech (LAUKKA; JUSLIN; BRESIN, 2005; PEREIRA, 2000; BARBOSA, 2009). Laukka, Juslin, & Bresin (2005) had listeners rate utterances conveying five acted emotions (anger, happiness, sadness, fear, and disgust) with respect to four emotional dimensions (activation, valence, potency, and intensity). The authors found that these emotions were associated with different patterns of judgments for these emotional dimensions, which suggests that these basic emotions can be described and distinguished by means of dimensions. In addition, some acoustic correlates were found in the utterances for all four dimensions. Laukka & Elfenbein (2012) have found that emotional dimensions related to the appraisal of the emotion-eliciting events (*e.g.* valence, novelty, urgency, goal conduciveness, etc.) can also be inferred reliably from vocal expressions, which suggests that speech can also signal information about the cognitive representation of events.

In spite of the great interest of speech researchers in the discrete and dimensional approaches, it is still unclear which of the two perspectives is more suitable to describe the expression of emotions in speech and there is a lack of studies which have directly compared these approaches in order to shed light on this question. Some authors have suggested that emotional dimensions may be better for distinguishing and describing the vocal expression of emotions than labels of discrete emotions, since some dimensions (*e.g.* activation) directly reflect the emotion-related changes in physiology (COWIE; CORNELIUS, 2003; SCHERER, 1986). In addition, emotions which have a similar pattern for some dimensions (*e.g.* activation and valence) share the same changes for some acoustic parameters (*e.g.* fundamental frequency and intensity) and this may cause confusion when trying to describe these emotions by means of discrete labels (LUGGER; YANG, 2007; PEREIRA, 2000). Emotional dimensions may be useful for research on speech synthesis and automatic recognition of emotions from speech, since their continuous quality allows the modeling of weaker emotional expressions, as well as gradual changes in speech expressiveness over time (GRIMM *et al.*, 2007; SCHRÖDER *et al.*, 2001; BARBOSA, 2009).

The present study was conducted to address these questions and thus contribute to the literature on emotional prosody. We carried out two perception experiments with utterances selected from spontaneous speech of a foreign language (Swedish). The use of utterances of an unknown foreign language is important to guarantee that listeners rely only on the prosody of the utterances to make their judgments (this is discussed in section 2.1). In the first experiment, Brazilian subjects rated on 5-point scales the expressivity of four basic emotions: joy, anger, sadness, and calmness. In the second, a distinct group of Brazilian subjects rated the expressivity of five emotional dimensions: activation, fairness, valence, motivation, and involvement. Our objectives were to investigate the relation between the perception of these basic emotions and emotional dimensions in speech, and to determine which acoustic cues among those investigated here are related to their perception. Based on the studies reported above, we hypothesize that each of these basic emotions is associated with a distinct pattern of perceptual judgments for these emotional dimensions and that the ratings for the discrete emotions and emotional dimensions correlate to some extent with some acoustic parameters extracted from the utterances.

2 Materials and methods

2.1 Authentic emotional speech

The corpus used in the present study consists of 40 speech samples from 5 Swedish female speakers,¹ each one with duration between 1 and 6 seconds and with acceptable quality for performing acoustic analysis. It was set up for a cross-cultural study on the perception of emotions through spontaneous speech which was carried out with Brazilian and Swedish listeners and showed no difference in the perception of the emotions between the subjects of both cultures, that is, the Brazilian listeners' perception of the emotions expressed by the speakers of this corpus was very similar to the native speakers' (SILVA; BARBOSA;

¹ The programs from which the utterances were extracted did not provide information regarding the age of the participants. According to our perception, the speakers were aged between 30 and 60 at the time of the recording.

ABELIN, 2016).² The reason for using this corpus is that the subjects who took part in the perception experiments reported here had no knowledge of the Swedish language. The use of utterances of an unknown foreign language prevents listeners from using words referring to emotions as clues to determine the emotional state of the speaker. Therefore, the listeners could rely only on the prosody and voice quality of the utterances to make their judgments.

The utterances of this corpus were extracted from spontaneous speech (talk shows and interviews) of the Swedish television and of one Swedish interview program which was freely available over the internet as podcasts. They were saved on the hard drive into wave sound format with a sampling frequency of 44.1 kHz (Mono).

2.2 Perception experiment I: basic emotions

2.2.1 Participants

13 subjects completed this experiment (7 women and 6 men). All of them were born and have lived most part of their lives in Brazil and have Portuguese as their mother language. They were either undergraduate or graduate students and reported having no hearing impairment. The average age of the judges was 28 years, ranging from 21 to 50 years. They also reported having no knowledge of the Swedish language.

2.2.2 Procedure

In this experiment, subjects were asked to rate on 5-point scales ranging from 0, “not at all adjective”, to 4, “very adjective”, the degree with which the speaker in each stimulus was expressing the discrete

² Studies comparing the perception of emotions through speech between speakers of Portuguese and of other languages are rare, but we can also mention the work by Peres (2014), which compared the perception of emotions expressed in Brazilian Portuguese utterances between Brazilian and English subjects. In this study, both groups recognized the emotions with better-than-chance rate, but the native listeners (Brazilian) performed better (90% against 66%).

emotions described by four adjectives: *joyful*, *angry*, *sad*, and *calm*.³ Therefore, the experiment consisted of 4 parts, which were completed in a single session. In each part the listeners evaluated one emotional adjective for all 40 speech samples. The stimuli were presented in a random order, but the adjectives were evaluated by all listeners in the order presented above. The experiment was developed in Portuguese and carried out over the internet through the “Survey Gizmo” online software (<http://www.surveygizmo.com/>). The link for accessing it was sent by email to the subjects who were interested in taking part. They were asked to use earphones and to do the experiment in a quiet room. One speech sample was presented on each screen along with its corresponding scale and it was reproduced automatically as the page was finished loading. After listening to the utterance, the subjects had to mark their response on the scale by clicking on the desired value and then click on the “next” button at the bottom of the page to proceed to the next stimulus. It was not possible to return to the previous page or to proceed to the next one without having marked the response on the scale.

2.3 Perception experiment II: emotional dimensions

2.3.1 Participants

This experiment was carried out with a group of judges different from that of experiment I. 20 subjects completed the experiment (7 men and 13 women). Their average age was 25 years, ranging from 18 to 34 years. They were either graduate or undergraduate students, were born and have lived most part of the life in Brazil and have Portuguese as their mother language. They reported having no knowledge of the Swedish language at all and no hearing impairment.

2.3.2 Procedure

In this experiment, judges were asked to rate on 5-point scales ranging from 0, “not at all adjective”, to 4, “very adjective”, the degree

³ The adjective “calm” is used here to balance the emotional space under investigation with an item that corresponds to an emotion of positive valence and low level of arousal. Thus, we have a representative label for each quadrant of the arousal-valence space (see, for example, James Russell’s circumplex model of affect (RUSSELL, 1980).

with which the speaker in each stimulus was expressing the emotional state described by emotional dimensions. The experiment consisted of 5 parts, which were carried out in a single session. In each part the listeners evaluated one emotional dimension for all 40 speech samples. The dimensions investigated in this experiment were: *activation* (“How agitated was the speaker?”), *fairness* (“How fair did the speaker consider what happened?”), *valence* (“How pleasant for the speaker was the situation he/she was in?”), *motivation* (“How motivated to act on the situation was the speaker?”), and *involvement* (“How involved is the speaker with the situation he/she was in?”). The stimuli were presented randomly, but the emotional dimensions were evaluated by all listeners in this order. After listening to each utterance, the subjects had to judge the degree of expressivity of the emotional dimension specific to that part of the experiment by clicking on the desired value of the 5-point scale and then click on the “next” button at the bottom of the page to listen to the next stimulus. The experiment was developed and run in Portuguese through the “Survey Gizmo” online software. The questions related to the dimensions (shown above) were presented along with the scales to guide the judges. The remaining of the procedure is the same as that followed in experiment I.

3 Analyses and results

To ensure comparable magnitudes with the normalized values of the acoustic parameters, the judges’ responses were linearly converted to a scale ranging from 0 to 1 (0, 0.25, 0.50, 0.75, and 1). Then, the 5-level responses were transformed into three categories: low (0 and 0.25), medium (0.50), and high (0.75 and 1). This was done to avoid the influence of outliers on the mean ratings for each utterance to be used in the following analyses, thus guaranteeing more reliable judgments. The statistical analyses reported in this paper were performed with the software R in its 3.1.2 version (R CORE TEAM, 2014).

3.1 Inter-rater reliability

The reliability of the listeners’ responses in both experiments was verified by means of the Fleiss’ kappa index (FLEISS, 1971), which is an index bounded between 0 and 1 (the closer to 1, the greater the agreement). The index was calculated separately for each adjective, considering the three categories of responses.

Table 1 shows the kappa figures for the discrete emotional labels and emotional dimensions as well as their corresponding z figure (this test is statistically significant for $\alpha = 0.001$ when $z > 3.09$). The agreement between the listeners' responses was statistically different from 0 for all adjectives ($p < 0.001$) and similar to other studies on the perception of emotions through speech (ALM; SPROAT, 2005; DEVILLERS *et al.*, 2006; BARBOSA, 2009).

TABLE 1
Kappa index for the discrete emotional labels and emotional dimensions and their corresponding z value

Experiment I		
Adjective	Kappa	z
<i>enraivecido</i> (angry)	$k = 0.465$	$z = 34.0$
<i>calmo</i> (calm)	$k = 0.408$	$z = 30.4$
<i>triste</i> (sad)	$k = 0.281$	$z = 21.0$
<i>alegre</i> (joyful)	$k = 0.231$	$z = 17.7$
Experiment II		
Adjective	Kappa	z
<i>agitado</i> (agitated)	$k = 0.359$	$z = 43.9$
<i>agradável</i> (pleasant)	$k = 0.280$	$z = 34.0$
<i>motivado</i> (motivated)	$k = 0.258$	$z = 31.1$
<i>justo</i> (fair)	$k = 0.199$	$z = 24.1$
<i>envolvido</i> (involved)	$k = 0.186$	$z = 22.4$

3.2 Comparing the perception of basic emotions and of emotional dimensions

In order to evaluate the relation between the perceptual judgments for the discrete emotions and emotional dimensions, correlation coefficients were computed between the judges' mean ratings for each utterance for the discrete emotions and for the emotional dimensions. The correlation coefficient measures the degree of association between two variables (DOWDY; WEARDEN; CHILKO, 2004).

Because some of the variables did not meet the normality condition for their distribution (according to the Shapiro–Wilk normality test), we used the Spearman's rank correlation (ρ), which is a nonparametric

alternative to the Pearson product-moment correlation. To perform this test the variables were transformed into ranks. Results are shown in Table 2. The Spearman's rank correlation coefficient is bounded between -1 and 1 , with 1 indicating perfect positive correlation (*i.e.* a given observation of one variable has the same rank order as the corresponding observation of the other variable) and -1 expressing perfect negative correlation (*i.e.* when a given observation of one variable has a high rank order, the corresponding observation of the other variable has a low rank order). The correlations were statistically significant (*i.e.* different from 0) for almost all discrete emotion – emotional dimension pairs.

The ratings for *joy* were positively correlated with fairness, valence, and motivation, which is in accordance with theoretical predictions which state that joy is an emotion of positive valence (ELLSWORTH; SCHERER, 2003; LAUKKA; ELFENBEIN, 2012; SCHERER, 1986). This indicates that according to the listeners' perception the speakers who were joyful were motivated to act on the situation and considered the situation pleasant and fair.

Anger was negatively correlated with fairness and valence and positively correlated with activation, motivation, and involvement. This means that there was a tendency for the utterances which were evaluated by the listeners with high level of anger to be also rated with high level for the dimensions of activation, motivation, and involvement (and thus with low level for the dimensions of fairness and valence). According to the listeners' perception, these speakers were very agitated, very involved and very motivated to act on the situation and did not consider the situation neither pleasant nor fair. This result is also in accord with the theoretical predictions and evidence presented in the literature for the relation of this basic emotion to valence, fairness, and activation (ELLSWORTH; SCHERER, 2003; LAUKKA; ELFENBEIN, 2012; SCHERER, 1986).

As expected, *sadness* was negatively correlated with activation, motivation, and involvement. Based on the literature, we also expected a negative correlation of this basic emotion with fairness and valence (since one expects that a person who is sad must have considered the event unfair and unpleasant). Although this was the case with the latter, the correlation with these two dimensions was not statistically significant.

Calmness was also negatively correlated with activation, motivation, and involvement. It differed from sadness for being positively correlated with fairness and valence, which was expected. In addition,

calmness was exactly the opposite of anger with regard to its correlation with the five emotional dimensions.

TABLE 2
Spearman's rank correlations between the listeners' judgments
(mean ratings for each utterance) for the discrete emotions
and for the emotional dimensions

	Activation	Fairness	Valence	Motivation	Involvement
Joy	0.17 ns	0.36*	0.56***	0.44**	0.23 ns
Anger	0.88***	-0.81***	-0.70***	0.69***	0.79***
Sadness	-0.49**	0.03 ns	-0.09 ns	-0.62***	-0.37*
Calmness	-0.94***	0.82***	0.69***	-0.72***	-0.82***

* p < 0.05 ** p < 0.01 *** p < 0.001 ns = not significant

It is also useful to analyze how the emotional dimensions of fairness, motivation, and involvement relate to the well-known dimensions of activation and valence according to the listeners' perception. The correlations between the basic emotions and the emotional dimensions presented in Table 2 suggest that fairness is more related to valence (*i.e.* they tended to be positively correlated with each other) whereas motivation and involvement seem to be more linked to activation. To better investigate this, the intercorrelations among the listeners' mean ratings for the emotional dimensions were also computed through the Spearman's rank correlation coefficient. Results are shown in Table 3. As expected, fairness presented a high and positive correlation coefficient with valence, whereas motivation and involvement presented high and positive correlations with activation and with each other.

This result indicates that the listeners' perception indeed separated fairness and valence from involvement, motivation, and activation. Put differently, there was a tendency for the utterances which were evaluated by the listeners with high values of the rating scales for the dimension of fairness to be also rated with high values for the dimension of valence. Conversely, the speech samples which were evaluated by the listeners with high values for activation tended to be also rated with high values for the dimensions of motivation and involvement.

TABLE 3
Intercorrelations (Spearman's rho) among the emotional dimensions assessed by the subjects in experiment II

	activation	fairness	valence	motivation	involvement
activation	—				
fairness	-0.77***	—			
valence	-0.58**	0.88***	—		
motivation	0.82***	-0.45*	-0.23 ns	—	
involvement	0.88***	-0.65***	-0.50**	0.82***	—

* $p < 0.01$ ** $p < 0.001$ *** $p < 10^{-5}$ ns = not significant

To examine to what extent the *linear combination* of the five emotional dimensions explains the variance of the listeners' judgments for the basic emotions, we performed a series of multiple linear regression models with basic emotion as the response variable and the listeners' judgments for the emotional dimensions as the predictor variables. However, because the emotional dimensions are highly intercorrelated (as shown in table 3), we applied principal component analysis on the emotional dimensions to eliminate covariate variables. This analysis showed that the first (PC1) and the second (PC2) principal components taken together account for **95%** of the total variance of the listeners' judgments for the five emotional dimensions (PC1 = 73%; PC2 = 22%). Therefore, the scores of the utterances in these principal components were taken as the predictor variables of the multiple linear regression models (representing the listeners' judgments for the five emotional dimensions). To correct for violations of normality and/or of constant variance (statistical assumptions for multiple linear regression analysis), we applied a log transformation to the response variable when necessary (for joy, calmness, and anger).

The model with *joy* as the response variable yielded an adjusted R^2 of **67%** ($F[2, 36] = 39.96, p < 10^{-09}$), which indicates that the linear combination of PC1 and PC2 explains a significant and relative high proportion of variance of the listeners' judgments for joy. There was no significant interaction between PC1 and PC2 in this model. The model with *calmness* as the response variable and with a significant interaction between PC1 and PC2 yielded an adjusted R^2 of **89%** ($F[3, 36] = 109.5, p < 10^{-15}$). The model with *anger* as the response variable

yielded an adjusted R^2 of **69%** ($F[2, 37] = 43.71, p < 10^{-09}$), and there was no significant interaction between PC1 and PC2. Finally, the model with *sadness* as the response variable and with a significant interaction between PC1 and PC2 yielded an adjusted R^2 of **68%** ($F[3, 36] = 28.16, p < 10^{-08}$).

3.3 Acoustic analysis

To investigate which acoustic parameters of speech correlate better with the listeners' judgments for the basic emotions and emotional dimensions and thus contribute to advance the knowledge of how these parameters vary as a function of the speaker's emotions, a set of acoustic features was automatically extracted from the utterances evaluated by the judges in the experiments by means of the script "Expression Evaluator", implemented for the software Praat (BOERSMA; WEENINK, 2011) by Barbosa (2009).⁴

This script analyzes the following acoustic features: fundamental frequency (f0), fundamental frequency first derivative (df0), global intensity, spectral tilt and Long-Term Average Spectrum (LTAS). All of these parameters have been reported in the literature as potential correlates of the vocal expression of emotions, since they may be affected by the psychophysiological responses triggered by the emotional processes (BARBOSA, 2009; FRICK, 1985; PITTAM; SCHERER, 1993; SCHERER, 1986). Fundamental frequency is an acoustic correlate of the rate of vocal fold vibration and is perceived as the pitch of the voice. Sound intensity, measured in decibels (dB), corresponds to the variations in the air pressure of a sound wave and it is the major contributor to the sensation of loudness of a sound. Spectral tilt, considered here as the difference of intensity between the 0 – 1250 Hz and 1250 – 4000 Hz frequency bands computed every ten points, measures the degree of the drop in intensity as the frequencies of the spectrum increase. LTAS is an intensity spectrum obtained from the average of several spectra extracted from the speech sample for a given frequency range. Various authors argue that the LTAS reduces the effect of individual linguistic segments on the spectral structure of speech, thus providing a spectral representation of the speaker's voice as a whole (PITTAM; GALLOIS;

⁴ This Praat script is available from the authors upon request.

CALLAN, 1990; SCHERER, 1982). The f_0 first derivative, computed as the difference in Hz between successive odd-numbered f_0 values, is used as a means of revealing abrupt changes in the intonation contour (BARBOSA, 2009).

In addition, spectral tilt and LTAS are acoustic correlates of vocal effort and voice quality, since the increase of vocal effort enhances the energy in the harmonics of high frequencies due to changes in subglottal pressure and in the characteristics of vocal fold vibration (LAUKKANEN *et al.*, 1997; TRAUNMÜLLER; ERIKSSON, 2000).

These acoustic parameters are computed by the script in terms of the following statistical descriptors (yielding a total of twelve parameters):

f_0 : median, interquartile semi-amplitude, skewness, and 99.5% quantile;

df_0 : mean, standard deviation, and skewness;

global intensity: skewness;

spectral tilt: mean, standard deviation, and skewness;

LTAS: slope (difference of mean intensity in dB between the bands 0 – 1000 Hz and 1000 – 4000 Hz of the LTAS).

The statistical descriptors related to f_0 and df_0 were normalized for inter-speaker variability through the z-score technique⁵ by using the following reference values (mean, standard deviation) of f_0 for adult females: (231 Hz, 120 Hz). Spectral tilt was normalized by dividing its value by the complete-band intensity median. The f_0 interquartile semi-amplitude is calculated as the difference between the 95% and 5% quantiles of f_0 , divided by two. It is, therefore, a measure analogous to f_0 range (f_0 maximum - f_0 minimum), but less sensitive to measurement errors, since it does not take into account the extreme values of the data. Similarly, the f_0 99.5% quantile is a measure analogous to f_0 maximum. Skewness indicates whether the distribution of the random variable (*i.e.* the measured values) is symmetric or asymmetric about its mean (and, in the latter case, whether the larger concentration of the values is on the left or on the right of the mean). The f_0 skewness is taken as the difference between f_0 mean and f_0 median, divided by the f_0 interquartile semi-amplitude.

⁵ The z-score is a statistical procedure that expresses the distance in terms of standard deviation units of an observation from its mean (here, the reference values).

The acoustic data were then correlated with the listeners' mean ratings for each utterance for the basic emotions (Table 4) and emotional dimensions (Table 5) by computing the Spearman's rank correlation. The parameters which correlated better with the listeners' judgments were f0 median, interquantile semi-amplitude, and 99.5% quantile, spectral tilt mean and standard deviation, and LTAS slope. In general, these parameters correlated better with the emotional dimensions. The parameters f0 median, f0 interquantile semi-amplitude, spectral tilt mean, and LTAS slope, for example, were significantly correlated with all five emotional dimensions, but only with some of the basic emotions. In addition, f0 skewness was significantly correlated only with fairness, motivation, and involvement.

Joy was not significantly correlated with any of the acoustic features. This may have occurred because of a possible lack of exemplars of this emotion in our corpus, or perhaps because the speakers of our corpus may not have expressed this emotion consistently through speech. The sign of the correlation coefficients (positive or negative) reveals that the perceived increase of *calmness* in the speakers' speech was associated with a decrease of f0 median, f0 interquantile semi-amplitude, f0 99.5% quantile, and of the energy concentrated in the harmonics of higher frequencies (indicated by spectral tilt mean and standard deviation, and LTAS slope).⁶ Higher ratings of *anger* were linked to an increase in f0 median, f0 interquantile semi-amplitude, f0 99.5% quantile, and in high-frequency energy. *Sadness* was associated with a decrease of f0 interquantile semi-amplitude and in high-frequency energy (indicated by spectral tilt standard deviation and LTAS slope) across the utterances.

⁶ Because spectral tilt and LTAS slope are estimated by the difference of intensity between the lower and higher frequency bands and the intensity drops as the frequencies of the spectrum increase, an increase in these parameters means *less* energy concentrated in the harmonics of higher frequencies due to a lower vocal effort used in the production of the utterance (TRAUNMÜLLER; ERIKSSON, 2000).

TABLE 4
Spearman's rank correlations between acoustic parameters and listeners' judgments (mean ratings for each utterance) for the basic emotions

Acoustic Parameters	Joy	Calmness	Anger	Sadness
Fundamental frequency (f0)				
median	0.13 ns	-0.65***	0.57***	-0.17 ns
interquantile semi-amplitude	0.13 ns	-0.56***	0.53***	-0.37*
99.5% quantile	0.10 ns	-0.33*	0.31*	-0.15 ns
skewness	-0.06 ns	0.26 ns	-0.26 ns	-0.01 ns
F0 first derivative				
mean	-0.18 ns	0.07 ns	0.03 ns	0.28 ns
standard deviation	0.17 ns	-0.22 ns	0.23 ns	-0.15 ns
skewness	-0.06 ns	0.00 ns	-0.02 ns	0.08 ns
Global intensity				
skewness	-0.18 ns	0.00 ns	0.09 ns	0.17 ns
Spectral tilt				
mean	-0.03 ns	0.54***	-0.54***	0.28 ns
standard deviation	-0.21 ns	0.41**	-0.35*	0.55***
skewness	0.10 ns	0.10 ns	-0.13 ns	-0.01 ns
Long-Term Average Spectrum - LTAS				
slope	0.08 ns	0.81***	-0.79***	0.38*

*p < 0.05 **p < 0.01 ***p < 0.001 ns = not significant

With respect to the emotional dimensions, the perceived increase in the degree of activation, motivation, and involvement was significantly linked to an increase in f0 median, f0 interquantile semi-amplitude and in high-frequency energy (reflected mainly in spectral tilt mean and LTAS slope). Activation was also significantly associated with an increase in f0 99.5% quantile, whereas motivation and involvement with a decrease in f0 skewness. Conversely, the perceived increase in fairness and valence was associated with a decrease in f0 median, f0 interquantile semi-amplitude and in high-frequency energy (also indicated by spectral tilt mean and LTAS slope). In addition, fairness was significantly linked to a decrease in f0 99.5% quantile and an increase in f0 skewness. It can

also be observed that the opposite behavior between the two groups of dimensions (which separate fairness and valence from involvement, motivation, and activation) observed in Table 2 and in Table 3 still holds for the correlations between these emotional dimensions and the acoustic parameters. This means that different patterns of changes in acoustic parameters characterize these groups.

TABLE 5

Spearman's rank correlations between acoustic parameters and listeners' judgments (mean ratings for each utterance) for the emotional dimensions

Acoustic Parameters	Activation	Fairness	Valence	Motivation	Involvement
Fundamental frequency (f0)					
median	0.75***	-0.57***	-0.38*	0.63***	0.72***
interquartile semi-amplitude	0.59***	-0.51***	-0.33*	0.48**	0.48**
99.5% quantile	0.40*	-0.34*	-0.19 ns	0.22 ns	0.29 ns
skewness	-0.29 ns	0.31*	0.25 ns	-0.31*	-0.44**
F0 first derivative					
mean	-0.14 ns	0.00 ns	-0.18 ns	-0.25 ns	-0.00 ns
standard deviation	0.28 ns	-0.20 ns	-0.09 ns	0.16 ns	0.20 ns
skewness	-0.06 ns	0.06 ns	0.03 ns	-0.13 ns	-0.09 ns
Global intensity					
skewness	-0.06 ns	-0.16 ns	-0.29 ns	-0.04 ns	0.09 ns
Spectral tilt					
mean	-0.59***	0.45**	0.38*	-0.54***	-0.63***
standard deviation	-0.47**	0.27 ns	0.13 ns	-0.47**	-0.28 ns
skewness	-0.00 ns	0.05 ns	0.13 ns	-0.02 ns	0.02 ns
Long-Term Average Spectrum - LTAS					
slope	-0.79***	0.72***	0.64***	-0.65***	-0.64***

*p < 0.05 **p < 0.01 ***p < 0.001 ns = not significant

4 Discussion

The present study was conducted to shed light on the relation between the perception of basic emotions and emotional dimensions in speech, as well as to identify some acoustic cues which may guide this process. For this purpose, a group of Brazilian subjects rated the expression of four basic emotions in utterances of a foreign language of which they had no knowledge (Swedish) and a separate group of Brazilian subjects rated the expression of five emotional dimensions for the same utterances. To further advance the knowledge of how emotions are expressed and perceived in everyday interactions, the corpus used in this study was composed of authentic emotional expressions as conveyed in spontaneous speech.

The novelty of this study lies in the direct comparison between the perception of some basic emotions and some emotional dimensions in speech, which provided evidence on how the two perspectives can be related to each other. The use of multiple linear regression analysis to assess this relation is also new, and proved effective. The majority of studies which suggested a relation between basic emotions and emotional dimensions have not done so in the form of a specific relationship, but only presented classes of basic emotions in a dimensional space. In summary, this work has empirically shown that the perception of basic emotions in speech can be described as a combination of emotional dimensions, which, in the study presented in this paper, tended to display a specific pattern for each basic emotion. In addition to contributing to the literature on emotion in general, this finding is also relevant to research on emotion and speech, as it provides researchers with empirical evidence which may help them to choose the best approach for their studies and to better interpret the results obtained with these approaches. Furthermore, three of the emotional dimensions investigated here (fairness, motivation, and involvement), despite being recognized by some theorists, have not been satisfactorily examined in studies on emotional prosody. Thus, this study also contributed by providing knowledge of other emotional dimensions, which can be used in expressive speech applications. The dimensions of motivation and involvement, for example, can be used in automatic processing of meetings, in order to detect heated arguments or periods of high excitement (WREDE; SHRIBERG, 2003) or in call center conversations (together with activation) to monitor the affective state of customers (VOGT; ANDRÉ; WAGNER, 2008).

Our results show that, apart from the classic dimensions of activation and valence, the dimensions of fairness, motivation, and involvement can also be inferred from speech. This is not only ensured by a significant agreement between the listeners' judgments for their expressivity (which was slight to fair but similar to other studies on the subject), but mainly because these dimensions were strongly and significantly correlated with some acoustic parameters, basic emotions and with activation and valence. This finding suggests that other emotional dimensions related to the appraisal of the eliciting event and to action readiness can also be inferred from speech. Investigating other emotional dimensions which could possibly be recognized from speech may advance our understanding of how emotions (and other affective phenomena) are expressed and perceived in speech, as well as of the expressive functions of speech prosody.

All the five emotional dimensions investigated here were significantly correlated to some extent with the basic emotions. In addition, the four basic emotions analyzed in the present study tended to have different patterns of perceptual judgments for these dimensions:⁷ *joy* (positive fairness, positive valence, and high motivation); *anger* (high activation, negative fairness, negative valence, high motivation, and high involvement); *sadness* (low activation, low motivation, and low involvement); *calmness* (low activation, positive fairness, positive valence, low motivation, and low involvement). As can be observed, with regard to activation and valence, these patterns were, in general, consistent with previous findings. In this way, this study contributed to better understand the underlying structure of these discrete emotions with regard to some emotional dimensions. Furthermore, the multiple linear regression analysis performed with each basic emotion and the two principal components corresponding to the listeners' judgments for the five emotional dimensions showed that indeed it is possible to describe the perception of basic emotions in speech as a combination of emotional dimensions, since the linear combination of the two principal components explained a significant and relative high (more than 50%) proportion of variance of the listeners' judgments for all the basic emotions (as revealed by the adjusted R^2 of the models).

⁷ Because its correlations with fairness and valence were not statistically significant, sadness could not be reliably distinguished from calmness in this respect.

The basic emotions and the emotional dimensions were also significantly correlated with some acoustic parameters extracted automatically from the utterances used as stimuli in the perception experiments, which indicates that the listeners relied partly on these acoustic features to judge the expressiveness of these emotions and dimensions. The most robust parameters were f_0 median, interquartile semi-amplitude, and 99.5% quantile, spectral tilt mean and standard deviation, and LTAS slope. These parameters confirm the relevance of fundamental frequency and voice quality in the communication of emotions through speech. Among the basic emotions, only anger and calmness correlated significantly with a considerable number of acoustic parameters. Joy was not significantly correlated with any of the acoustic features and sadness correlated moderately with three of them. All five emotional dimensions, on the other hand, correlated significantly with various acoustic parameters. Therefore, it is possible to conclude that, in general, the acoustic parameters correlated better with the emotional dimensions. This result is consistent with some previous studies, which have suggested that emotional dimensions are more suitable to distinguish and describe the vocal expression of emotions than labels of discrete emotions (COWIE; CORNELIUS, 2003; LUGGER; YANG, 2007; PEREIRA, 2000; BARBOSA, 2009). Research on the automatic recognition of emotions from speech and speech synthesis can benefit from this finding, since the use of emotional dimensions may allow the reliable identification and synthesis of more subtle expressions of emotions and changes in speech expressiveness over time (GRIMM, *et al.*, 2007; SCHRÖDER *et al.*, 2001; BARBOSA, 2009).

Acknowledgments

This work was funded by a fellowship from the São Paulo Research Foundation – FAPESP (2012/04254-4 and 2013/06082-9) to the first author. We thank all the subjects who took part in the perception experiments reported here.

References

ALM, C. O.; SPROAT, R. Perceptions of emotions in expressive storytelling. In: INTERSPEECH 2005, 2005, Lisboa. *Proceedings of Interspeech 2005*. Lisboa: ISCA Archive, 2005. p. 533-536. Disponível em: <http://www.isca-speech.org/archive/interspeech_2005>. Acesso em: 14 jul. 2016.

AMIR, N.; MIXDORFF, H.; OFER AMIR, D. R.; DIAMOND, G. M.; PFITZINGER, H. R.; LEVI-ISSERLISH, T.; ABRAMSON, S. Unresolved Anger: Prosodic analysis and classification of speech from a therapeutical setting. In: INTERNATIONAL CONFERENCE ON SPEECH PROSODY 2010, 5., 2010, Chicago. *Proceedings...* Chicago, 2010. Disponível em: <<http://speechprosody2010.illinois.edu/program.php>>. Acesso em: 14 jul. 2016.

BANSE, R.; SCHERER, K. R. Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, v. 70, n. 3, p. 614-636, 1996. doi.org/10.1037/0022-3514.70.3.614

BARBOSA, P. A. Detecting changes in speech expressiveness in participants of a radio program. In: INTERSPEECH 2009, 2009, Brighton. *Proceedings of Interspeech 2009 - Speech and Intelligence*. Londres: Causal Productions, 2009. p. 2155-2158. Disponível em: <http://www.isca-speech.org/archive/interspeech_2009/i09_2155.html>. Acesso em: 18 nov. 2016.

BARBOSA, P. A. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 20, n. 1, p. 11-27, 2012. Disponível em: <<http://periodicos.letras.ufmg.br/index.php/relin/article/view/2571>>. Acesso em: 18 nov. 2016.

BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer (Versão 5.2.25)*. Disponível em: <<http://www.praat.org>>. Acesso em: 20 Mai. 2011.

BROSCH, T; MOORS, A. Valence. In: SANDER, D.; SCHERER, K.R. (Org.). *The Oxford Companion to Emotion and the Affective Sciences*. Oxford: Oxford University Press, 2009. p. 401-402.

CORNELIUS, R. R. Theoretical approaches to emotion. In: ISCA TUTORIAL AND RESEARCH WORKSHOP (ITRW) ON SPEECH AND EMOTION, 2000, Newcastle. *Proceedings...* Disponível em: <http://www.isca-speech.org/archive_open/speech_emotion/>. Acesso em: 14 jul. 2016.

COWIE, R.; CORNELIUS, R. R. Describing the emotional states that are expressed in speech. *Speech communication*, v. 40, n. 1, p. 5-32, 2003. doi.org/10.1016/S0167-6393(02)00071-7

DARWIN, C. *A expressão das emoções no homem e nos animais*. São Paulo: Companhia das Letras, 2009 [1872].

DEVILLERS, L.; COWIE, R.; MARTIN, J.-C.; DOUGLAS-COWIE, E.; ABRILIAN, S.; MCRORIE, M. Real life emotions in French and English TV video clips: An integrated annotation protocol combining continuous and discrete approaches. In: INTERNATIONAL CONFERENCE ON LANGUAGE RESOURCES AND EVALUATION, Fifth., 2006, Genoa. *Proceedings...* Disponível em: <<http://www.lrec-conf.org/proceedings/lrec2006/>>. Acesso em: 14 jul. 2016.

DOWDY, S.; WEARDEN, S.; CHILKO, D. *Statistics for research*. 3. ed. Hoboken: John Wiley & Sons, 2004. doi.org/10.1002/0471477435

EKMAN, P. An argument for basic emotions. *Cognition & Emotion*, v. 6, n. 3/4, p. 169-200, 1992. doi.org/10.1080/02699939208411068

ELLSWORTH, P. C.; SCHERER, K. R. Appraisal processes in emotion. In: DAVIDSON, R. J.; SCHERER, K. R.; GOLDSMITH, H. H. (Org.). *Handbook of affective sciences*. New York: Oxford University Press, 2003. p. 572-595.

FLEISS, J. L. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, v. 76, n. 5, p. 378-382, 1971. doi.org/10.1037/h0031619

FONTAINE, J. Dimensional Emotion Models. In: SANDER, D.; SCHERER, K.R. (Org.). *The Oxford Companion to Emotion and the Affective Sciences*. Oxford: Oxford University Press, 2009. p. 119-120.

FOWLES, D. C. Arousal. In: SANDER, D.; SCHERER, K. R. (Org.). *The Oxford Companion to Emotion and the Affective Sciences*. Oxford: Oxford University Press, 2009. p. 50-50.

FRICK, R. W. Communicating emotion: The role of prosodic features. *Psychological Bulletin*, v. 97, n. 3, p. 412-429, 1985. doi.org/10.1037/0033-2909.97.3.412

FRIJDA, N. H.; MARKAM, S.; SATO, K.; WIERS, R. Emotions and emotion words. In: RUSSELL, J. A.; FERNANDEZ-DOLS, J. M.; MANSTEAD, A. S. R.; WELLENKAMP, J. C. (Org.) *Everyday Conceptions of Emotion: An Introduction to the Psychology, Anthropology and Linguistics of Emotion*. NATO ASI series D: Behavioural and social sciences. Nova Iorque: Kluwer Academic/Plenum Publishers, 1995. v. 81, p. 121-143. doi.org/10.1007/978-94-015-8484-5_7

GRIMM, M.; KROSCHEL, K.; MOWER, E.; NARAYANAN, S. Primitives-based evaluation and estimation of emotions in speech. *Speech Communication*, v. 49, n. 10, p. 787-800, 2007. doi.org/10.1016/j.specom.2007.01.010

JOHNSTONE, T.; SCHERER, K. R. Vocal communication of emotion. In: LEWIS, M.; HAVILAND, J. M. (Org.) *Handbook of emotions*. 2. ed. Nova Iorque: Guilford, 2000. p. 220-235.

KEHREIN, R. The prosody of authentic emotions. In: SPEECH PROSODY 2002, 2002, Aix-en-Provence. *Proceedings of Speech Prosody 2002*. Disponível em: <<http://sprosig.isle.illinois.edu/sp2002/papers.htm>>. Acesso em: 14 jul. 2016.

LAUKKA, P.; ELFENBEIN, H. A. Emotion appraisal dimensions can be inferred from vocal expressions. *Social Psychological and Personality Science*, v. 3, n. 5, p. 529-536, 2012. doi.org/10.1177/1948550611428011

LAUKKA, P.; JUSLIN, P.; BRESIN, R. A dimensional approach to vocal expression of emotion. *Cognition & Emotion*, v. 19, n. 5, p. 633-653, 2005. doi.org/10.1080/02699930441000445

LAUKKANEN, A-M; VILKMAN, E.; ALKU, P.; OKSANEN, H. On the perception of emotions in speech: the role of voice quality. *Logopedics Phoniatrics Vocology*, v. 22, n. 4, p. 157-168, 1997. doi.org/10.3109/14015439709075330

LUGGER, M., YANG, B. An incremental analysis of different feature groups in speaker independent emotion recognition. In: INTERNATIONAL CONGRESS OF PHONETIC SCIENCES, 16., 2007, Saarbrücken. *Proceedings....* Disponível em: <<http://www.icphs2007.de/>>. Acesso em: 14 jul. 2016.

MATSUMOTO, D.; EKMAN, P. Basic emotions. In: SANDER, D.; SCHERER, K.R. (Org.). *The Oxford Companion to Emotion and the Affective Sciences*. Oxford: Oxford University Press, 2009. p. 69-72.

MAUSS, I. B.; ROBINSON, M. D. Measures of emotion: A review. *Cognition and Emotion*, v. 23, n. 2, p. 209-237, 2009. doi.org/10.1080/02699930802204677

MESQUITA, B.; FRIJDA, N. H. Cultural Variations in Emotions: A Review. *Psychological Bulletin*, v. 112, n. 2, p. 179-204, 1992. doi.org/10.1037/0033-2909.112.2.179

PEREIRA, C. Dimensions of emotional meaning in speech. In: ISCA TUTORIAL AND RESEARCH WORKSHOP (ITRW) ON SPEECH AND EMOTION, 2000, Newcastle. *Proceedings of the ISCA Workshop on Speech and Emotion*. Disponível em: <http://www.isca-speech.org/archive_open/speech_emotion/>. Acesso em: 14 jul. 2016.

PERES, D. O. A manifestação da emoção na fala: estudo perceptual com falantes nativos e não nativos. *Estudos Linguísticos*, São Paulo, v. 43, n. 1, p. 10-21, 2014.

PITTAM, J.; GALLOIS, C.; CALLAN, V. The long-term spectrum and perceived emotion. *Speech Communication*, Amsterdam, v. 9, n. 3, p. 177-187, 1990. doi.org/10.1016/0167-6393(90)90055-E

PITTAM, J.; SCHERER, K. R. Vocal expression and communication of emotion. In: LEWIS, M.; HAVILAND, J. M. (Org.). *Handbook of emotions*. New York: Guilford Press, 1993. p. 185-198.

R CORE TEAM. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>>. Acesso em: 5 fev. 2014.

RUSSELL, J. A. A circumplex model of affect. *Journal of Personality and Social Psychology*, v. 39, p. 1161-1178, 1980. doi.org/10.1037/h0077714

SCHERER, K. R. Methods of research on vocal communication: Paradigms and parameters. In: SCHERER, K. R.; EKMAN, P. (Org.). *Handbook of methods in nonverbal behavior research*. Cambridge: Cambridge University Press, 1982. p. 136-198.

SCHERER, K. R. Vocal affect expression: a review and a model for future research. *Psychological Bulletin*, v. 99, n. 2, p. 143-165, 1986. doi.org/10.1037/0033-2909.99.2.143

SCHERER, K. R. Emotion. In: HEWSTONE, M.; STROEBE, W. (Org.). *Introduction to Social Psychology: A European perspective*. 3. ed. Oxford: Blackwell, 2000. p. 151-191.

SCHERER, K. R. Vocal communication of emotion: a review of research paradigms. *Speech Communication*, v. 40, n. 2, p. 227-256, 2003. doi.org/10.1016/S0167-6393(02)00084-5

SCHERER, K. R.; BANSE, R.; WALLBOTT, H. G. Emotion Inferences from Vocal Expression Correlate across Languages and Cultures. *Journal of Cross-Cultural Psychology*, Sage Publications, v. 32, n. 1, p. 76-92, 2001. doi.org/10.1177/0022022101032001009

SCHERER, K. R.; ELLGRING, H. Multimodal expression of emotion: Affect programs or componential appraisal patterns? *Emotion*, v. 7, n. 1, p. 158-171, 2007. doi.org/10.1037/1528-3542.7.1.158

SCHLOSBERG, H. A scale for the judgement of facial expressions. *J. Exp. Psychology*, Association for the Advancement of Affective Computing., v. 29, p. 497-510, 1941.

SCHLOSBERG, H. Three dimensions of emotion. *Psychological Review*, American Psychological Association, v. 61, n. 2, p. 81-88, 1954. doi.org/10.1037/h0054570

SCHRÖDER, M.; COWIE, R.; DOUGLAS-COWIE, E.; WESTERDIJK, M.; GIELEN, S. C. Acoustic correlates of emotion dimensions in view of speech synthesis. In: INTERSPEECH 2001, 2001, Aalborg. *Proceedings of Interspeech 2001*. Disponível em: <http://www.isca-speech.org/archive/eurospeech_2001/index.html>. Acesso em: 14 jul. 2016.

SILVA, W.; BARBOSA, P. A.; ABELIN, Å. Cross-cultural and cross-linguistic perception of authentic emotions through speech: An acoustic-phonetic study with Brazilian and Swedish listeners. *DELTA*, São Paulo, v. 32, n. 2, p. 449-480, 2016. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44502016000200449&lng=en&nrm=iso>. Acesso em: 18 nov. 2016.

SMITH, C. A.; ELLSWORTH, P. C. Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, v. 48, n. 4, p. 813-838, 1985. doi.org/10.1037/0022-3514.48.4.813

TRAUNMÜLLER, H; ERIKSSON, A. Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America*, v. 107, n. 6, p. 3438-3451, 2000. doi.org/10.1121/1.429414

VOGT, T.; ANDRÉ, E.; WAGNER, J. Automatic Recognition of Emotions from Speech: A Review of the Literature and Recommendations for Practical Realisation. In: PETER, C.; BEALE, R. (Org.). *Affect and Emotion in Human-Computer Interaction: From Theory to Applications (Lecture Notes in Computer Science)*. Heidelberg, Germany: Springer, 2008. v. 4868, p. 75-91. doi.org/10.1007/978-3-540-85099-1_7

WREDE, B.; SHRIBERG, E. Spotting “hot spots” in meetings: human judgments and prosodic cues. In: INTERSPEECH 2003, 2003, Genebra. *Proceedings of Interspeech 2003*. Genebra: ISCA Archive, 2003. p. 2805-2808. Disponível em: <http://www.isca-speech.org/archive/eurospeech_2003/e03_2805.html>. Acesso em: 19 out. 2016.

Prosódia de enunciados declarativos e interrogativos totais nas variedades de Salvador, Fortaleza e Rio de Janeiro

Prosody of declarative and interrogative modalities in Salvador, Fortaleza and Rio de Janeiro varieties

Leticia Rebollo Couto

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
rebollocouto@yahoo.fr

Carolina Gomes da Silva

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil
Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
carolinagsufpb@gmail.com

Luma da Silva Miranda

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
CAPES
lumah.miranda@gmail.com

Resumo: Este trabalho tem como objetivo descrever diferenças entre variedades dialetais do português falado em Salvador, Fortaleza e Rio de Janeiro, Brasil, tanto nas sentenças declarativas quanto nas interrogativas. A análise de parâmetros prosódicos (frequência fundamental, duração, intensidade) é feita no núcleo final de sentenças oxítonas, paroxítonas e proparoxítonas. Há diferenças prosódicas sistemáticas entre as duas modalidades (declarativa e interrogativa): um tom alto ou subida de F0 no núcleo das interrogativas e, ao contrário, um tom baixo ou descida de F0 no núcleo das declarativas. Variações dialetais também são observadas. Os acentos tonais nucleares $H + L * L\%$ para as declarativas e $L + H * L\%$ para as interrogativas, propostos por Moraes (2008), são observados nos falantes do Rio de Janeiro – sendo este último truncado ou comprimido

em sentenças oxítonas, devido à falta do material postônico. Entretanto, nas variedades de Salvador e Fortaleza, as interrogativas também podem ser realizadas com um padrão L + H*H%. Maior elisão de sílabas é observada nos falantes de Fortaleza.

Palavras-chave: prosódia; entoação; sentenças declarativas; sentenças interrogativas totais; variação dialetal.

Abstract: This work aims at describing differences between the dialectal varieties of Portuguese spoken at Salvador, Fortaleza and Rio de Janeiro, Brazil, in both the declarative and interrogative modalities. Analyses of prosodic parameters (fundamental frequency, duration, intensity) are made on the final nucleus of oxytonic, paroxytonic and proparoxytonic sentences. There are systematic prosodic changes between modalities – with a higher / rising F0 in interrogative nucleus where it is lower / decreasing for declaratives. The interrogative nucleus are also longer and stronger than the declaratives. Dialectal variations are also observed. The accentual patterns H + L*L% for declaratives and L + H*L% for interrogatives proposed by Moraes (2008) are observed in Rio de Janeiro speakers – the latter being truncated or compressed in oxytonic sentences due to lack of post-tonic material. Meanwhile, in the Salvador and Fortaleza varieties, interrogatives may also be performed with a L + H*H% pattern. More syllable elisions are also observed in Fortaleza speakers.

Keywords: prosody; intonation; declarative sentences; yes/no questions; dialectal variation.

Recebido em: 5 de janeiro de 2017.

Aprovado em: 14 de março de 2017.

1 Introdução

A proposta deste trabalho é descrever e comparar elementos prosódicos de enunciados interrogativos totais e declarativos do *corpus* experimental AMPER (Atlas Multimédia Prosódico do Espaço

Românico)¹ para o português do Brasil: projeto AMPER Por (Brasil). Os dados que analisamos neste artigo são referentes a três pontos de coleta, duas capitais na região nordeste do Brasil – Salvador e Fortaleza – e uma na região sudeste – Rio de Janeiro. Analisamos a variação do padrão entonacional e a implementação dos parâmetros acústicos frequência fundamental (F0), duração e intensidade no núcleo e no pré-núcleo dos seguintes enunciados: “O Renato gosta do Renato”, “O pássaro gosta do pássaro”, “O bisavô gosta do bisavô”, nas suas modalidades declarativas e interrogativas totais (respostas sim/não).

Partindo de estudos já realizados acerca do português do Brasil (doravante PB), e em particular sobre as variedades de Salvador, Fortaleza e Rio de Janeiro, pretendemos com este estudo discutir resultados dos falares destas três cidades, e seus padrões melódicos de declarativas e interrogativas. Um dos principais problemas dos estudos comparativos é a disparidade de escolhas metodológicas.

Os dados do AMPER são dados de fala controlada, experimental ou de laboratório. O estudo das funções prosódicas de segmentação e de hierarquização da fala se faz sobre frases construídas a fim de que variem sistematicamente suas estruturas morfossintáticas, o que permite a oposição de um traço particular entre duas frases, que tenham estruturas iguais (RILLIARD, 2011). Esta função de demarcação e de hierarquização da fala é uma hipótese clássica sobre as funções da prosódia que se encontra descrita em Rossi (1987), Reis (1995), Vaissière (1997) ou Ladd (2008) e nós nos afiliamos a esta abordagem teórica. Os dados que coletamos e analisamos para o AMPER-Por (Brasil) nas três capitais (Rio de Janeiro, Fortaleza e Salvador) podem ser considerados como dados de fala lida ou fala atuada, de acordo com as diferentes condições de produção e oralização dos enunciados, e de acordo com as diferentes conceitualizações teóricas de fala lida, atuada ou espontânea.

¹ O projeto AMPER (Atlas Multimédia Prosodique de l’Espace Roman – Atlas Multimédia Prosódico do Espaço Românico) é coordenado por Michel Contini do Centro de Dialectologia da Universidade de Grenoble, na França e visa contemplar línguas românicas, tais como o italiano, o francês, o castelhano e o galego, além das variedades do português europeu (PE) e do português brasileiro (PB). A pesquisa relativa à variação prosódica do Português é coordenada pela Lurdes de Castro Moutinho, do Centro de Investigação de Línguas e Culturas da Universidade de Aveiro. AMPER-POR disponível em: <http://pfonetica.web.ua.pt/AMPER-POR.htm>.

Segundo Gomes da Silva (2014, p. 55-56), a primeira característica de um texto espontâneo, tal como sugere Blanche Benveniste (1998), é que este nunca é uma produção acabada e que na sua oralização deixa ver suas etapas de produção. A segunda característica é que na fala espontânea aparecem comentários sobre a escolha do léxico e a dificuldade de encontrar a palavra adequada. A terceira característica é que, nesse tipo de discurso, as tentativas lexicais são conservadas, uma vez que não é possível apagar o que se acaba de dizer. A fala espontânea pode apresentar titubeios, repetições, falsos começos. A fala espontânea além de não apresentar uma sequência linear, sintagmática, de uma seleção em ausência, está constituída por relações paradigmáticas, expostas, em presença. O falante vai e volta, sobre sintagmas já apresentados, seja para corrigi-los ou completá-los, acrescentando comentários ou aspectos que considere relevantes, de uma forma que poderíamos definir como um modo *online*, ou “sobre la marcha”, como foi traduzido em espanhol. Trata-se de uma produção simultânea à sua oralização, no que diverge da fala lida ou atuada.

Ao contrário da fala espontânea, na fala lida ou atuada, seja experimental ou não, o texto já está pronto e é oralizado a posteriori. Na fala atuada, o texto é aprendido de cor e praticado a fim de conseguir uma *performance* de interação, com efeitos expressivos controlados. Na fala lida o texto está pronto e é oralizado a posteriori, mas diferentemente da fala atuada o texto não é aprendido de cor. As fronteiras podem ser sempre tênues e dependem dos gêneros e eventos de interação. No caso dos dados do AMPER, o falante tem o texto pronto, mesmo que não seja com palavras, mas com imagens sob os olhos, portanto estaríamos mais inclinados a considerá-los como dados de leitura, são frases muito controladas quanto à sua estrutura sintática e fonética que são por sua vez oralizadas sem contexto de interação. Neste estudo comparamos durante nossa análise dados do AMPER aos do projeto ALIB (Atlas Linguístico do Brasil)² para as três capitais. No caso dos dados do ALIB, o falante é apresentado a

² O Projeto “Atlas Linguístico do Brasil” foi criado em 1996, sob a coordenação de um Comitê Nacional, presidido pela Professora Doutora Suzana Cardoso, da Universidade Federal da Bahia. O ALIB, mais especificamente, se constitui pela recolha de dados em 250 localidades, tanto nas capitais dos estados como outras cidades de grande e médio porte, linguisticamente representativas. Com relação aos informantes, foram gravados um total de 1100 informantes. Para a coleta dos dados, o Projeto ALIB conta com um inquérito composto por três questionários: (i) Questionário Morfosintático (QMS); (ii) Questionário Semântico-Lexical (QSL) e (iii) Questionário Fonético-Fonológico

uma situação, a um contexto no qual ele tem que inserir a frase esperada ou proposta com certa intencionalidade. Consideramos que no ALIB os dados coletados são de fala atuada.³ Os falantes do ALIB seguem um questionário que inclui questões para a apuração de diferenças prosódicas relativas à natureza das frases. Como exemplo desses contextos de interação, apresentamos aqui os contextos de coleta das interrogativas. São apresentados 4 contextos, com instruções e respostas esperadas para as interrogativas, os dois primeiros para interrogativas disjuntivas e os dois últimos para interrogativas totais (CUNHA, 2005):

1) *Ó, meu amigo, você prefere vinho ou cerveja?* (**Resposta Esperada**)

Se você/o(a) senhor(a) quer oferecer uma bebida a um amigo e quer saber se ele prefere vinho ou cerveja, como é que você/o(a) senhor(a) se dirige a ele e pergunta? (**Instrução**)

2) *Ó, meu amigo, você toma leite ou café?* (**Resposta Esperada**)

Se você/o(a) senhor(a) quer saber se o seu amigo toma leite ou café, como é que você/o(a) senhor(a) se dirige a ele e pergunta? (**Instrução**)

3) *Você vai sair hoje?* (**Resposta Esperada**)

Se você/o(a) senhor(a) quer saber se alguém vai sair hoje, como é que você/o(a) senhor(a) pergunta? (**Instrução**)

4) *Eu vou sair hoje, doutor?* (**Resposta Esperada**)

Uma pessoa está internada em um hospital e quer saber do médico se vai sair naquele dia. Como é que pergunta? (**Instrução**)

(QFF), que inclui questões para apuração de diferenças prosódicas, relativas à natureza das frases, que podem ser interrogativas, afirmativas ou imperativas.

³ Tanto nos dados do AMPER quanto nos dados do ALIB, há “surpresas” na coleta de dados. Mesmo que o texto esteja pronto, no AMPER há falantes que não produzem os artigos ou que mudam as estruturas, razão pela qual se coletam seis de cada enunciado. Em alguns casos, itens lexicais inesperados podem aparecer. Imagens são interpretáveis, mas só são consideradas e selecionadas as frases que reproduzam o modelo ou mais se aproximem em número de sílabas ao modelo esperado. Também no ALIB, Cunha (2005) observa que não se produzem as frases esperadas, quase nunca, na interação de fala atuada.

Para as declarativas, são apresentados 3 contextos diferentes, com instruções e respostas esperadas, o primeiro para declarativa neutra e os dois últimos para declarativas com marcas de expressividade (CUNHA, 2006). No entanto, Silvestre (2012, p. 59) sinaliza que, além de ter apenas uma questão concernente à declarativa neutra, os informantes não produziram de maneira satisfatória as informações desejadas e, ao contrário do esperado, não houve a possibilidade de comparar frases idênticas ou, ao menos, terminadas pelo mesmo vocábulo. Por essa razão, em seu trabalho, Silvestre (2012) compara enunciados declarativos neutros produzidos ao longo dos demais questionários.

1) *Você vai sair hoje.* (**Resposta Esperada**)

Uma pessoa está internada num hospital e quer saber do médico se vai sair naquele dia. Se a resposta for positiva, como é que o médico responde? (**Instrução**)

2) *Oh, gente, estou muito aborrecido com o que aconteceu.* (**Resposta Esperada, expressando desagrado, rispidez**)

Você / o(a) senhor(a) quer dizer a algumas pessoas que estão presentes que você / o(a) senhor(a) está muito aborrecido com o que aconteceu. Como é que você / o(a) senhor(a) diz? (**Instrução**)

3) *Oh, gente, estou muito feliz com o resultado do trabalho.* (**Resposta Esperada, expressando contentamento, polidez**)

Você / o(a) senhor(a) quer dizer a algumas pessoas que estão presentes que você / o(a) senhor(a) está muito feliz com o resultado do trabalho. Como é que você / o(a) senhor(a) diz? (**Instrução**)

As estruturas comparadas são próximas e têm em geral o mesmo número de sílabas, tanto no projeto AMPER quanto no ALIB, que utiliza dados de entrevista e enunciados paroxítonos. Entretanto, consideramos que os dados do ALIB são de fala atuada os dados do AMPER são de fala lida, sendo ambos experimentais.

Esta comparação da variação prosódica de enunciados do AMPER Por (Brasil), referentes a três capitais brasileiras, tem quatro objetivos: 1) caracterizar o padrão declarativo e interrogativo total de Salvador, Fortaleza e Rio de Janeiro verificando se os acentos tonais H + L* L% para as sílabas nucleares das declarativas e L + H* L% para as sílabas nucleares das interrogativas, propostos por Moraes (2008) se realizam

com estes dados do projeto AMPER;⁴ 2) discutir os resultados referentes às duas capitais, Salvador e Fortaleza, de acordo com o trabalho de Lira (2009) que analisou dados dessas capitais; 3) comparar os resultados do AMPER aos resultados da análise de dados do projeto ALIB propostos por Silva (2011) e Silvestre (2012) para as três capitais; 4) descrever a variação da sílaba tônica com relação às sílabas pretônica e postônica nucleares em termos de duração, de frequência fundamental e de intensidade, comparando convergências e divergências de implementação destes parâmetros acústicos nas três variedades locais analisadas. Para a realização deste estudo, seguimos os padrões comparativos propostos por Moraes (2008) para o estado do Rio de Janeiro, os de Antunes (2012) para o estado de Minas Gerais, os de Nunes (2015) para o estado de Santa Catarina e os de Seara e Rebollo-Couto (2011) para dados comparados de Santa Catarina e Rio de Janeiro, considerando ainda a discussão proposta por Lira (2009) com relação à variação entonacional em cinco capitais da região nordeste.

2 Pesquisas anteriores sobre os enunciados declarativos e interrogativos totais

Para as descrições de contornos declarativos e interrogativos totais, Moraes (2008) propõe dois acentos tonais nucleares contrastantes. A inflexão final apresenta um movimento final descendente (\downarrow) nas modalidades assertivas, exclamativas e interrogativas parciais, ou ainda com final circunflexo (\wedge), ou seja, descendente-ascendente-descendente, na modalidade interrogativa total. Seguindo a notação métrica autossegmental, Moraes (2008) propõe enunciados declarativos neutros com um padrão melódico $H+L^*L\%$, ou seja, uma queda entre as sílabas finais pretônica e tônica, seguido de um nivelamento na postônica. Para os enunciados interrogativos totais neutros, propõe a notação $L+H^*L\%$, com diferenças de implementação segundo o alinhamento do pico na tônica, tardio para as perguntas e adiantado para os pedidos. Vale ressaltar que, segundo Cunha (2000), esses acentos podem variar de acordo com

⁴ De acordo com a teoria métrica autossegmental a notação (*) corresponde à sílaba tônica lexical e a notação (%) ao tom de fronteira do enunciado. O núcleo é a última sílaba tônica do enunciado, e o pré-núcleo tudo o que está à sua esquerda. O núcleo é portanto o final do enunciado, marcado pelo último acento lexical (*) e o tom de fronteira (%).

a área dialetal a que pertence o falante. O par contrastante de acentos tonais (*pitch accent*) para o PB, entre as declarativas H+ L* L% e as interrogativas totais neutras L + H* L% propostos por Moraes (2008), tem sido sistematicamente encontrado em diversos estudos entonacionais e em diversos pontos do país.

Particularmente com relação às descrições do Rio de Janeiro, nos interessa destacar os resultados obtidos por Silva (2011) e Silvestre (2012), trabalhando com os dados do projeto ALIB. Para Silvestre (2012, p. 82-83):

Rio de Janeiro. A entoação nos enunciados assertivos neutros cariocas, já amplamente descritos no PB, apresentou em nosso corpus o mesmo comportamento outrora evidenciado por Moraes, Cunha, Tenani e outros autores: tom baixo nas sílabas que compõem o acento pré-nuclear e tons semelhantemente baixos observados até a última sílaba pretônica de I, na qual a F0 alcança o seu pico e posteriormente apresenta o movimento descendente final. O enunciado “O sol tá nascendo”, dito pela falante carioca da primeira faixa etária, jovem, dá, mais uma vez, amostra do padrão. Nele, a F0 de 157Hz na primeira sílaba tônica de I ascende 14% até alcançar o pico de 179Hz na última sílaba pretônica e decrescer 22% até a fronteira final.

E Silva (2011, p. 102-103) encontra duas variantes do mesmo padrão para as interrogativas totais na capital do Rio de Janeiro:

A questão total da capital fluminense apresentou, em geral, a seguinte configuração inicial: pico inicial hospedado na primeira sílaba tônica, seguido de declinação ao longo do interior da frase. A configuração circunflexa final apresentou dois tipos de comportamento na tônica final: pico melódico [+ antecipado] seguido de movimento descendente e movimento ascendente seguido de pico melódico [+ atrasado]. Ambos alcançam valores que superam os obtidos pelo primeiro pico. Os enunciados “A pessoa que tá internada vai sair hoje?” e “A minha alta vai ser hoje?”, apresentados a seguir, foram produzidos, respectivamente, pelo homem e pela mulher da primeira faixa etária. O contorno inicial comum a ambos apresenta pico inicial na primeira sílaba tônica, seguidos de movimento descendente que no primeiro enunciado,

apresenta variação de 36 Hz e na segunda, de 21 Hz. A diferença entre essas frases é relativa à manifestação do pico da última sílaba tônica. Na primeira frase, observa-se um movimento de subida cujo ápice é atingido à direita dessa sílaba. Na segunda frase, por outro lado, o pico ocorre no início da tônica, sendo sucedido por um movimento descendente.

Ou seja, os trabalhos de Silva (2011) e Silvestre (2012) confirmam a proposta de Moraes (2008) para o Rio de Janeiro. E no caso de Salvador e Fortaleza? No caso dessas duas capitais do Nordeste contamos com os resultados de Lira (2009), cujos dados são neste artigo reanalisados em função da comparação com o Rio de Janeiro, considerando, além do contorno melódico analisado na tese de Lira (2009), variações de F0, duração e intensidade.

Lira (2009) analisa apenas os enunciados interrogativos totais e parciais, de Salvador e Fortaleza. No que diz respeito aos contornos melódicos das questões totais nessas duas capitais, a autora propõe dois padrões contrastantes para as duas variedades regionais. A proposta é de uma **tônica final alta** seguida de postônicas baixas para **Fortaleza**, e inversamente, de uma **tônica final baixa**, seguida de postônica alta em **Salvador**, movimento ascendente que se inicia já na sílaba tônica (LIRA, 2009, p. 106-107)

Essa **tônica final baixa** para **Salvador**, mas com movimento já ascendente e pico nas postônicas é exemplificada, nos três padrões acentuais, da seguinte forma para os interrogativos totais de Salvador (LIRA, 2009, p. 101-106):

- a) O final **oxítono** termina em movimento ascendente “bisavô” com pequeno apêndice sobre a tônica final (2009, p. 104).
- b) O final **paroxítono**, “Renato”, com a perda da última sílaba, também termina como o padrão oxítono com uma subida melódica moderada na sílaba tônica, de 98 a 129 Hz considerando os valores médios de realização de um falante (2009, p. 104).
- c) E o final **proparoxítono**, “pássaro”, termina ora com um movimento ascendente sobre a tônica final que se prolonga sobre as postônicas finais, ora com uma sílaba tônica final que se situa em um nível alto, mas seguido de uma queda sobre a postônica final (2009, p. 106).

Essa **tônica final alta** para **Fortaleza** é exemplificada nos três padrões acentuais, da seguinte forma para os interrogativos totais de Fortaleza (LIRA, 2009, 9 p. 4-96), sendo os padrões oxítono e paroxítono bem próximos em oposição ao proparoxítono:

- a) O final **oxítono**, “bisavô”, termina com uma subida melódica mais nítida na sílaba tônica, de 96 a 124 Hz, considerando os valores médios de realização de um falante (2009, p. 95).
- b) O final **paroxítono**, “Renato”, com a perda da última sílaba, também termina com uma subida melódica moderada na sílaba tônica, de 96 a 105 Hz considerando os valores médios de realização de um falante (2009, p. 94).
- c) E, o final **proparoxítono**, “pássaro”, termina com uma subida melódica após a tônica final, que passa de 124 a 130 Hz e chega a 180 Hz na última postônica final, considerando os valores médios de realização de um falante, sem supressão das últimas sílabas átonas, e com pico na última átona (2009, p. 96)

Silva (2011), em pesquisa utilizando o *corpus* do projeto ALIB, analisa os enunciados interrogativos com final paroxítono (“O meu amigo vai ter alta hoje?”; “Você vai sair hoje?” e “A minha alta vai ser hoje?”).

Para Salvador, Silva (2011, p. 93-94) encontra dois padrões, um padrão de **tônica alta seguida de postônica baixa**, que corresponde ao acento proposto por Moraes (2008) L + H* L% mas também um acento com um tom de fronteira alto L + H* H%, que foi o padrão encontrado por Lira (2009), **tônica alta seguida de postônica alta**:

O traçado inicial mais comum encontrado nas frases analisadas de Salvador apresenta nível alto na tônica, que também pode ser manifestado na pretônica, seguido de movimento descendente que se alonga até a pretônica final. Em relação ao contorno do final da frase, foram encontrados dois comportamentos: uma configuração circunflexa formada por F0 baixa na pretônica, subida melódica com pico alinhado à direita na tônica e descida na postônica; e um movimento ascendente final, hospedado nas sílabas tônica e postônica. Nos dados analisados de Salvador, o segundo pico apresenta-se mais proeminente que o primeiro.

O enunciado “Você vai sair hoje?”, [vo 'se 'vaj sa 'i 'o' ʒi], produzido pela falante mulher da segunda faixa etária, apresenta o padrão que foi mais recorrente nos dados analisados de Salvador. Nessa frase, o nível mais alto da frase recai sobre a primeira tônica. Um movimento descendente de 54 Hz ocupa o interior do enunciado. Na sílaba tônica, a frequência aumenta 50 Hz em direção ao final da sílaba, onde alcança seu pico, decrescendo 34 Hz na postônica.

No enunciado “Eu estou de alta?”, [ew iʃ 'to dʒi 'aw tɐ], realizado pela falante mulher da primeira faixa etária, observa-se o padrão melódico menos recorrente nos dados analisados da capital baiana. Essa frase possui pico melódico na pretônica, seguido de movimento descendente de 25 Hz até a pretônica final e subida de 38 Hz da F0 na última tônica, que permanece alta na postônica.

Para Fortaleza, Silva (2011, p. 84-85) encontra um **padrão de tônica alta** seguido de postônica baixa, que corresponde ao acento proposto por Moraes (2008) $L + H^* L\%$, com ensurdecimento sistemático da última sílaba postônica, potencialmente ligado à escolha da unidade lexical “hoje”:

Nos enunciados coletados em Fortaleza, foi observado um único padrão de questão total. A primeira configuração da frase apresenta movimento descendente que se espalha desde a primeira até a última pretônica do enunciado. A segunda configuração, por sua vez, caracteriza-se por movimento ascendente-descendente no interior da tônica final, cujo pico possui traço [+ atrasado]. O comportamento da postônica não pôde ser descrito devido ao seu ensurdecimento em todos os dados.

No enunciado “O meu amigo vai ter alta hoje?”, [u 'mew ã 'mi gu vaj te 'aw tɐ 'o ʒi], dito pela falante mulher da primeira faixa etária, constata-se que a proeminência de 216 Hz do início da frase localiza-se na primeira sílaba pretônica, sendo seguida por um decréscimo de 52 Hz ao longo das sílabas do interior da frase. A configuração circunflexa final está hospedada na última sílaba tônica, com pico alinhado à sua direita. Essa configuração é menos destacada que a primeira, apresentando variação de 38 Hz na subida e 19 Hz na descida da F0. (SILVA, 2011, p. 84-85)

Silvestre (2012, p. viii), em pesquisa utilizando o mesmo *corpus* do projeto ALIB, apresenta para todos os enunciados declarativos (“Você vai estar de alta hoje.”; “Eu dou alta pra ele.” e “O sol está nascendo”) e nas três localidades, Salvador, Fortaleza e Rio de Janeiro, um contorno descendente no núcleo, com notação fonológica H+L*L%, sendo que a diferença entre os contornos do Rio de Janeiro e os do nordeste estariam no ataque, ou pré-núcleo dos enunciados declarativos:

Ao padrão observado nas capitais nortistas e nordestinas, atribui-se a mesma notação que Cunha (2005) propôs para a asserção neutra em Recife, com base no corpus do projeto NURC: H*___H+L*L%. Para o padrão majoritariamente observado nas capitais do centro-oeste e do sudeste do Brasil, atribui-se notação semelhante a de Cunha (2005) e Moraes (2008) para a asserção neutra, ambos com base na fala carioca: L+H*___H+L*L%. (SILVESTRE, 2012, p. viii)

Os trabalhos de Silva (2011) e Silvestre (2012) utilizam o *corpus* do ALIB para a análise prosódica dos enunciados interrogativos e declarativos, respectivamente, com o objetivo de caracterizar as diferenças entonacionais entre as capitais brasileiras.

3 Procedimentos metodológicos

Para a feitura desse trabalho, adotamos a metodologia empregada pelo Projeto AMPER (MOUTINHO *et alii*, 2011). O *corpus* compreende a gravação de uma série de sentenças declarativas e interrogativas totais neutras geradas através de estímulos visuais. Detalhamos a seguir por partes: o *corpus de pesquisa*, os *falantes da pesquisa* e *apresentação das etapas de análise*.

3.1 O *corpus* da pesquisa

O corpus desta análise está composto por 6 enunciados, formados por 3 enunciados declarativos, aqueles que afirmam a realidade ou a possibilidade de um fato, e 3 interrogativos totais, aqueles que correspondem ao caráter afirmativo ou negativo da predicação, ou seja, respondem-se com sim ou não. Os enunciados são analisados a partir de três repetições de cada, chegando a um total de 18 enunciados analisados

para cada falante. Todos eles possuem 10 sílabas contemplando final proparoxítono, paroxítono e oxítono. Estes enunciados foram obtidos por meio de gravações de três frases, repetidas três vezes por cada falante. A metodologia do AMPER prevê a gravação de 6 repetições dos dados para assim se justificar a escolha das três melhores. As frases são as seguintes, de acordo com seus núcleos paroxítono, paroxítono e oxítono:

- ◆O pássaro gosta do pássaro. ◆O pássaro gosta do pássaro?
- ◆O Renato gosta do Renato. ◆O Renato gosta do Renato?
- ◆O bisavô gosta do bisavô. ◆O bisavô gosta do bisavô?

Escolhemos três frases de cada modalidade produzida por cada falante. Nossos critérios foram: qualidade da gravação (algumas gravações apresentaram problemas técnicos, como por exemplo, enunciado truncado) e com a relação sinal-ruído alta o suficiente para realizar as análises acústicas deste trabalho.

Os programas de análise do AMPER geram uma curva de F0 média, deslexicalizada, a partir das três realizações escolhidas para cada enunciado por falante. A técnica de deslexicalização (RILLIARD, 2014, p. 30), baseada num processo de ressíntese da curva melódica, sem conteúdo segmental, assegura uma transmissão eficaz da informação prosódica normalizando eventuais efeitos microprosódicos e fonotáticos. A reiteração conserva uma estrutura silábica clara, e, portanto, todas as informações relacionadas à variação rítmica regional. A segmentação foi feita no Praat e a análise com o Matlab. O procedimento está detalhado em Romano *et alii* (2011).

Os dados foram coletados com a metodologia do projeto AMPER, com os enunciados adaptados para o PB. A gravação dos enunciados se coleta a partir da leitura de imagens. Os falantes recebem estímulos visuais para produzi-los, propostos por Moraes e Abraçado (2005).

Figura 1 – Exemplos de enunciado (A) declarativo e (B) interrogativo a ser produzido pelos falantes



A: O pássaro gosta do pássaro.

B: O pássaro gosta do pássaro?

Uma vez que o texto está pronto, e é oralizado a posteriori, após um período de treinamento e repetição, consideramos que estes dados coletados estão mais próximos aos dados de leitura e não aos dados de fala atuada, como os do projeto ALIB para as capitais brasileiras, coletados a partir de um contexto de interação e com maior controle da expressividade. Em ambos os casos estamos diante de fala experimental (ou de laboratório), gerada a partir de um elemento motivador, mas o fato de que o texto não esteja terminado, e não tenha sido produzido simultaneamente ao seu processo de oralização, nos faz considerar nossos dados como mais próximos aos dados de leitura (o texto está sob os olhos do falante) e não aos de fala atuada (ele já sabe o texto de cor, repetiu e praticou bastante a sua elocução).

3.2 Falantes da pesquisa

Os falantes desta pesquisa são originários de três cidades: Salvador e Fortaleza, na região nordeste do Brasil, e Rio de Janeiro, na região sudeste. Todos os 6 falantes, um homem e uma mulher de cada localidade, têm pelo menos nível de escolaridade médio concluído, são todos da categoria “jovens” de 20 a 35 anos. Os falantes do Rio de Janeiro estão no último ano de graduação em Letras, preparando seu TCC, o que pode influenciar no padrão de leitura que encontramos bem marcado para os falantes desta cidade com relação aos das duas capitais do nordeste que não têm nível superior.

Os falares das cidades de Salvador, Fortaleza e Rio de Janeiro são três “sotaques” facilmente reconhecíveis e estereotipados ou difundidos na mídia em todo o Brasil, embora faltem testes de percepção que atestem as condições de reconhecimento desses falares (CUNHA, 2006). Nesse sentido destacamos as pesquisas que começam a ser realizadas na

UFRJ, sob a orientação de João Antônio de Moraes para a descrição de movimentos melódicos perceptivamente relevantes (MIRANDA, 2015, p. 19) do Português do Brasil.⁵

3.3 Etapas de análise

A fase de análise de dados para esta pesquisa foi a segmentação dos enunciados no programa computacional PRAAT (BOERSMA; WEENINK, 2014), a fim de observar o comportamento das vogais tônicas, pretônicas e postônicas dos enunciados, como se implementaram os parâmetros acústicos F0, duração e intensidade de cada uma das vogais dos enunciados produzidos. Posteriormente, geramos gráficos dos parâmetros acústicos: frequência fundamental (F0), intensidade e duração vocálica através de scripts do próprio PRAAT e do programa Interface_AMPER_beta11 (RILLIARD; LAI, 2008), de acordo com as convenções apresentadas em Contini *et alii* (2009) ou em Romano *et alii* (2011). A partir desses gráficos fizemos a descrição fonética e a análise fonológica do contorno entonacional dos enunciados. No caso de vogais que não se realizam, a convenção é marcar um valor de F0 igual a 50 Hz o que faz que a curva pareça descontínua nos traçados e gráficos gerados.

Do ponto de vista fonético, para a realização deste estudo levamos em consideração o valor da duração e da intensidade das vogais do núcleo de cada enunciado (última sílaba acentuada) e da frequência fundamental (F0). A medida de F0 é feita em três pontos de cada vogal, começo meio e fim, introduzindo assim uma normalização temporal comparável à normalização apresentada por Xu (2004) e Arantes (2015). Para obter valores estáveis de cada um destes parâmetros, optamos por considerar a média de três repetições de cada vogal, de acordo com a metodologia AMPER. As variações de intensidade, ligadas às condições de gravação (em particular o nível da gravação e a distância entre o falante e o microfone) são aqui negligenciáveis, uma vez que consideramos, com a metodologia AMPER, as diferenças de nível médio de intensidade entre vogais próximas; essas diferenças são pouco influenciáveis pelas condições de gravação.

⁵ São trabalhos que destacam a entoação modal, através de estudos fonéticos experimentais da entoação do português do Brasil, a partir do método de ressíntese e estilização da curva de F0 em uma abordagem perceptiva baseada no modelo holandês IPO (MIRANDA, 2015).

A duração das vogais é considerada e não a das sílabas ou a das unidades V-V, o que constitui uma aproximação discutível do ritmo da fala (BARBOSA, 2007), mas as restrições inerentes ao corpus AMPER tornam impossível uma normalização estrita das influências intrínsecas dos fonemas; essas influências são controladas pela utilização das mesmas frases (e, portanto, dos mesmos fonemas), para todos os falantes. Sendo assim, neste trabalho observamos, com esta metodologia, padrões de alongamento e de redução da duração vocálica.

Analisamos o comportamento da F0 das vogais a partir do contraste entre os valores em Hz da vogal da primeira sílaba tônica e das vogais pretônicas e postônicas de cada enunciado. Do ponto de vista fonológico levamos em consideração o contorno melódico dos enunciados, tais como foram descritos inicialmente por Moraes (2008) e confirmados pelos trabalhos de Lira (2009), Silva (2011) e Silvestre (2012).

4 Resultados e discussões

A partir dos gráficos de F0, duração e intensidade, desenvolvemos nossa análise sobre os acentos tonais nucleares dos enunciados de cada localidade, a partir dos três padrões acentuais do português: começamos com o padrão nuclear paroxítono, o mais frequente e produtivo (1 sílaba tônica e 1 átona final), seguido do padrão proparoxítono, o mais longo (1 sílaba tônica e 2 sílabas átonas finais) e do padrão oxítono, o mais polêmico por aparecer truncado ou comprimido segundo as diferentes interpretações fonológicas (1 sílaba final tônica).

4.1 Resultados e discussões: núcleos paroxítonos

A categoria acentual paroxítona é a categoria mais frequente do português em termos de produtividade, apresentando, segundo Cintra (1997), uma frequência de distribuição de aproximadamente 70% em termos de distribuição de padrões acentuais nos vocábulos em português. O padrão acentual das paroxítonas é analisado a partir dos três parâmetros acústicos: F0, duração e intensidade em enunciados declarativos e interrogativos totais.

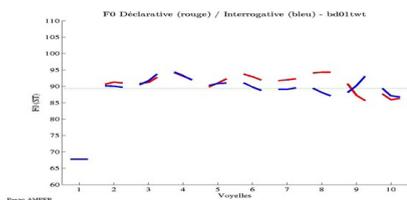
Em termos de F0, na variedade de **Salvador**, tanto o homem quanto a mulher realizam nas frases declarativas (em vermelho) um padrão ascendente e alto na pretônica do núcleo, “Renato”, sílaba 8,

descendente na sílaba tônica, sílaba 9, e baixo na sílaba postônica final, sílaba 10.

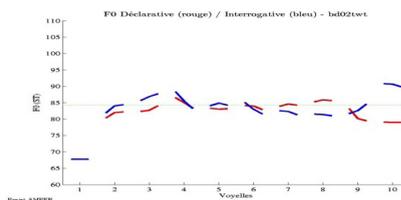
Os enunciados interrogativos (em azul) apresentam uma descida da F0 na sílaba pretônica, sílaba 8, uma curva ascendente na tônica, sílaba 9, e um comportamento variável na postônica, baixo/ descendente para a mulher e alto/ ascendente para o homem.

Figura 2 – Variedade de Salvador: escolaridade média

Enunciado 1: O Renato gosta do Renato./ O Renato gosta do Renato?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Salvador



B: F0 do homem de Salvador

As vogais elididas são representadas, por convenção, nos gráficos AMPER com 50 Hz para a F0 e com 0 (zero) duração e intensidade, portanto, a barra baixa na primeira vogal significa queda do artigo inicial.

No pré-núcleo “Renato”, ver sílabas 2, 3 e 4, temos o mesmo comportamento na declarativa e na interrogativa, embora nas interrogativas a F0 seja mais alta no homem: um movimento ascendente que culmina na postônica, uma tônica ascendente e uma postônica alta descendente. Trata-se de um comportamento ascendente progressivo, sem a queda abrupta da F0 na sílaba tônica da declarativa, como se observa no núcleo, sílaba 9, correspondendo à notação proposta por Silvestre (2012), L + H* como característica dos falares do norte e nordeste. Do ponto de vista fonológico, as duas realizações confirmam o padrão nuclear H + L* L% para as declarativas (vermelho), mas só a mulher confirma o padrão nuclear L + H* L% para as interrogativas (azul), proposto por Moraes (2008). No homem, o padrão realizado corresponde a L + H* H%, já registrado por Lira (2009) e Silva (2011) para Salvador.

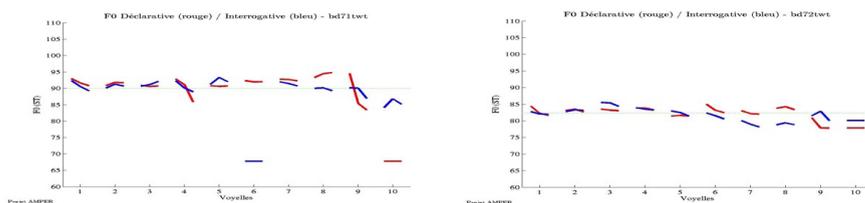
Na variedade de **Fortaleza**, tanto o homem quanto a mulher realizam nas frases declarativas (em vermelho) um padrão ascendente e alto na pretônica do núcleo, “Renato”, sílaba 8, descendente na sílaba

tônica, sílaba 9, e baixo na sílaba postônica final, sílaba 10, elidida pela mulher.

Os enunciados interrogativos (em azul) apresentam uma F0 baixa na sílaba pretônica, sílaba 8, uma curva ascendente-descendente na tônica, sílaba 9, e uma queda de F0 na sílaba postônica final, tanto para a mulher quanto para o homem.

Figura 3 – Variedade de Fortaleza: escolaridade média

Enunciado 1: O Renato gosta do Renato./ O Renato gosta do Renato?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Fortaleza

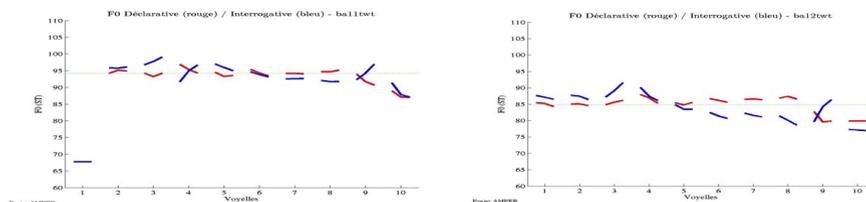
B: F0 do homem de Fortaleza

No pré-núcleo “Renato”, ver sílabas 2, 3 e 4, temos o mesmo comportamento na declarativa e na interrogativa: um movimento ascendente que culmina na postônica. Trata-se de um comportamento alto e suavemente ascendente progressivo, sem a queda abrupta da F0 na sílaba tônica da declarativa, como se observa no núcleo, sílaba 9, não correspondendo à notação proposta por Silvestre (2012), L + H*, como característica dos falares do norte e nordeste, mas sim à notação H*. Do ponto de vista fonológico, as duas realizações confirmam o padrão nuclear H + L* L% para as declarativas (vermelho), mas só o homem confirma o padrão nuclear L + H* L% para as interrogativas (azul), proposto por Moraes (2008). Na mulher, o padrão realizado corresponde a H + H* L% já registrado por Lira (2009) e Silva (2011) para Fortaleza.

Na variedade de **Rio de Janeiro**, tanto o homem quanto a mulher realizam nas frases declarativas (vermelho) um padrão ascendente de F0 na pretônica, do núcleo, “Renato”, descendente na sílaba tônica e baixo na sílaba postônica que não é elidida.

Figura 4 – Variedade Rio de Janeiro: escolaridade média

Enunciado 1: O Renato gosta do Renato./ O Renato gosta do Renato?
 Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Rio de Janeiro

B: F0 do homem de Rio de Janeiro

No pré-núcleo “Renato”, o comportamento do padrão é bem diferente do comportamento do núcleo, ver sílabas 2, 3 e 4, corresponde ao acento tonal H*, proposto para Rio de Janeiro por Moraes (2008) e Silvestre (2012). Trata-se de um comportamento ascendente progressivo com pico na postônica nas declarativas, sem a queda abrupta de F0 do núcleo, ver sílabas 8, 9 e 10. Do ponto de vista fonológico as duas realizações confirmam o padrão H + L* L% para as declarativas (vermelho) e o padrão L + H* L% para as interrogativas (azul), proposto por Moraes (2008).

Em termos de duração, a distribuição das durações por sílaba é muito mais regular, com menos diferença entre vogais acentuadas e não acentuadas ou átonas nos dados do Rio de Janeiro, do que nos de Salvador e Fortaleza, para as declarativas (vermelho) e para as interrogativas (azul).

Figura 5 – Duração de enunciado com núcleo paroxítono: mulheres

Enunciado 1: O Renato gosta do Renato./ O Renato gosta do Renato?
 Declarativas (vermelho) e interrogativas (azul).
 Duração das mulheres de Salvador, Fortaleza, e Rio de Janeiro⁶

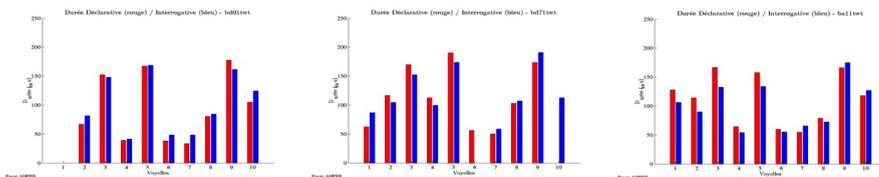
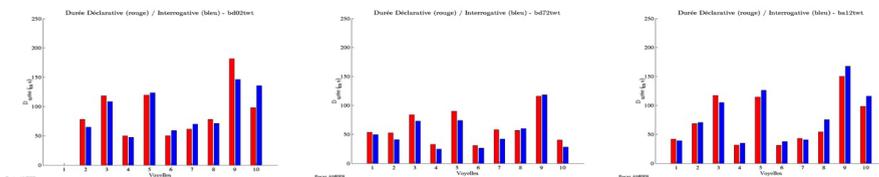


Figura 6 – Duração de enunciado com núcleo paroxítono: homens

Enunciado 1: O Renato gosta do Renato./ O Renato gosta do Renato?
 Declarativas (vermelho) e interrogativas (azul).
 Duração dos homens de Salvador, Fortaleza, e Rio de Janeiro



As vogais das sílabas tônicas de “Renato”, sílaba 3 no pré-núcleo, “gosta”, sílaba 5, e “Renato”, sílaba 9 no núcleo do enunciado, são mais longas do que as vogais das sílabas átonas que as sucedem em Fortaleza como no Rio de Janeiro.⁷

No Rio de Janeiro, as sílabas postônicas, 4, 6 e 10 são sempre realizadas, o que pode ser atribuído a um padrão escolar de leitura, aprendido e praticado nos anos de formação e exposição à variante

⁶ As vogais elididas são representadas, por convenção, nos gráficos AMPER com 50 Hz para a F0 e com 0 (zero) duração e intensidade. A seleção de frases se faz em função dessa elisão, quando ocorrer, são escolhidas três frases que tenham a vogal elidida no mesmo lugar, caso seja impossível encontrar três frases com os mesmos pontos de elisão selecionam-se apenas duas.

⁷ Para uma comparação mais precisa teria sido necessária a normalização da duração das vogais, entretanto, as vogais tônicas são tão mais longas que as átonas, que acreditamos que tal procedimento não deve alterar este padrão.

de leitura frente à fala espontânea. Esse padrão se repete nos dados de Fortaleza, apesar do ensurdecimento e das duas supressões de centros silábicos das átonas, sílaba 6 na declarativa “gosta” e sílaba 10 na declarativa “Renato”. Em termos de duração, o padrão mais acentual, com maior diferença entre sílabas átonas e tônicas, é o de Salvador. Trata-se de uma descrição qualitativa e não quantitativa dessa diferença que precisaria ser sistematicamente analisada em trabalhos futuros.

O falante masculino do Rio de Janeiro divide seu enunciado simetricamente em dois grupos rítmicos, com dois picos de duração nas duas tônicas no pré-núcleo, “Renato” e “gosta”, sílabas 3 e 5, respectivamente, e com dois picos nas duas últimas sílabas do núcleo, “Renato”, sílabas 9 e 10, com alongamento da postônica final, como estratégia rítmica de leitura (na nossa interpretação).

Em termos de intensidade, nas três localidades, a intensidade da sílaba tônica no núcleo é mais importante na interrogativa (azul), sílaba 9, do que na declarativa (vermelho). O comportamento da mulher do Rio de Janeiro se diferencia das demais, neste caso, pela queda abrupta progressiva de intensidade na sílaba átona final do enunciado, sílaba 10, procedimento estratégico característico da leitura que não se verifica nas outras duas falantes.

Figura 7 – Intensidade de enunciado com núcleo paroxítono: mulheres

Enunciado 1: O Renato gosta do Renato./ O Renato gosta do Renato?
 Declarativas (vermelho) e interrogativas (azul).
 Intensidade das mulheres de Salvador, Fortaleza, e Rio de Janeiro

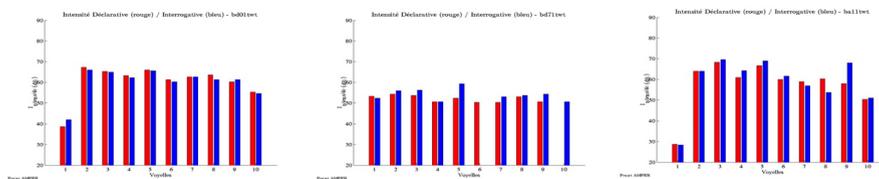
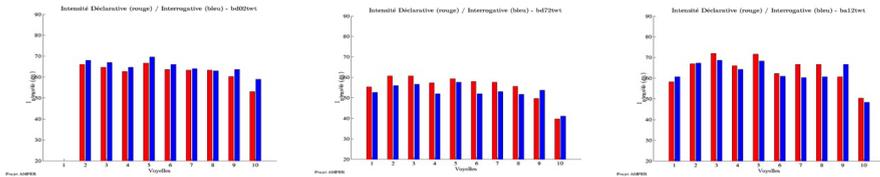


Figura 8 – Intensidade de enunciado com núcleo paroxítono: homens

Enunciado 1: O Renato gosta do Renato./ O Renato gosta do Renato?
Declarativas (vermelho) e interrogativas (azul).

Intensidade dos homens de Salvador, Fortaleza, e Rio de Janeiro



O comportamento do homem do Rio de Janeiro, mais escolarizado, se diferencia das demais, neste caso, pela queda abrupta progressiva de intensidade na sílaba átona final do enunciado, sílaba 10, procedimento estratégico característico da leitura que não se verifica nos outros dois falantes masculinos.

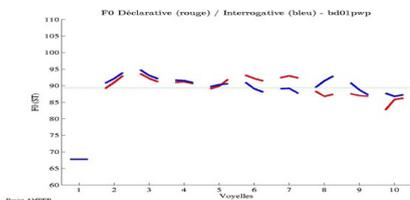
4.2 Resultados e discussões: núcleos proparoxítonos

A categoria acentual proparoxítona é a categoria menos frequente do português em termos de produtividade (REIS, 1995). Considera-se que seja um núcleo mais longo que se estende por três sílabas, quando a última sílaba não for elidida, transformando-se nesse caso num padrão paroxítono. O padrão acentual das proparoxítonas é analisado a partir dos três parâmetros acústicos: F0, duração e intensidade, em enunciados declarativos e interrogativos totais.

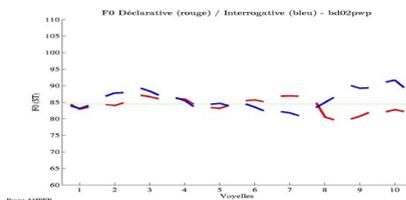
Em termos de F0, na variedade de **Salvador**, nem o homem nem a mulher elidem sílabas finais. Tanto o homem quanto a mulher realizam nas frases declarativas um padrão alto na pretônica, do núcleo, “pássaro”, mas descendente na sílaba tônica e baixo na sílaba postônica.

Figura 9 – Variedade de Salvador: escolaridade média

Enunciado 2: O pássaro gosta do pássaro./ O pássaro gosta do pássaro?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Salvador



B: F0 do homem de Salvador

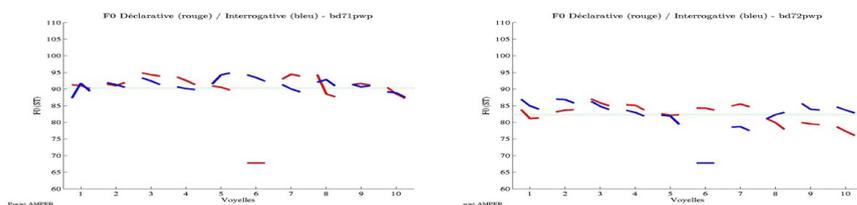
No pré-núcleo “pássaro”, o comportamento do padrão é bem diferente do comportamento no núcleo, ver sílabas 2, 3 e 4, para o pré-núcleo. Trata-se de um comportamento alto ou ascendente na sílaba tônica, sílaba 2, sem a queda abrupta de F0 do núcleo, sílaba 8 nas declarativas, tendo um padrão L + H* para o pré-núcleo e um padrão H + L* L% para o núcleo das declarativas. Nas interrogativas se realizam os dois padrões previstos para Salvador, um com final baixo, L + H* L% para a mulher e L + H* H% para o homem.

Do ponto de vista fonológico as duas realizações confirmam o padrão H + L* L% para as declarativas (vermelho) e o padrão L + H* L% para as interrogativas (azul), proposto por Moraes (2008), sendo que no homem se observa a variante descrita por Lira (2009) e Silva (2011) com final alto para Salvador.

Na variedade de **Fortaleza**, nem o homem nem a mulher elidem sílabas finais. Tanto o homem quanto a mulher realizam nas frases declarativas (vermelho) um padrão alto na pretônica, do núcleo, “pássaro”, alto, mas descendente na sílaba tônica e baixo na sílaba postônica.

Figura 10 – Variedade de Fortaleza: escolaridade média

Enunciado 2: O pássaro gosta do pássaro./ O pássaro gosta do pássaro?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Fortaleza

B: F0 do homem de Fortaleza

No pré-núcleo “pássaro”, o comportamento do padrão é bem diferente do comportamento no núcleo, ver sílabas 2, 3 e 4, para o pré-núcleo. Trata-se de um comportamento alto ou ascendente na sílaba tônica, sílaba 2, sem a queda abrupta de F0 do núcleo, sílaba 8 nas declarativas, tendo um padrão L + H* para o pré-núcleo e um padrão H + L* L% para o núcleo das declarativas. Nas interrogativas (azul), se realizam dois padrões para Fortaleza, um com final baixo, L + H* L% para a mulher e L + H* H% para o homem.

Do ponto de vista fonológico, as duas realizações confirmam o padrão H + L* L% para as declarativas (vermelho) e o padrão L + H* L% para as interrogativas (azul), proposto por Moraes (2008), sendo que no homem também se observa na variedade de Fortaleza a variante descrita por Lira (2009) e Silva (2011) com final alto para Salvador.

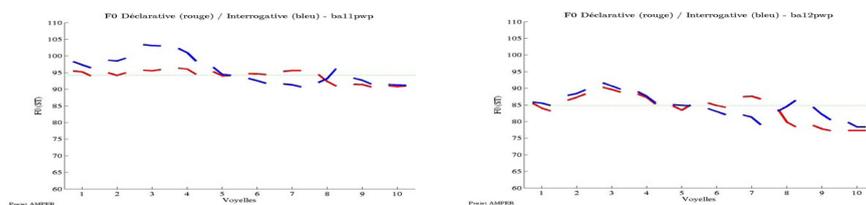
Ainda sob o ponto de vista fonológico, as duas realizações confirmam o padrão H + L* L% para as declarativas (vermelho) e o padrão L + H* L% para as interrogativas (azul), proposto por Moraes (2008). No pré-núcleo “pássaro”, o comportamento do padrão é bem diferente do comportamento no núcleo, ver sílabas 2, 3 e 4, para o pré-núcleo. Trata-se de um comportamento alto ou ascendente na sílaba tônica, sílaba 2, sem a queda abrupta de F0 do núcleo, sílaba 8 nas declarativas, tendo um padrão L + H* para o pré-núcleo e um padrão H + L* L% para o núcleo das declarativas. Nas interrogativas se realizam os dois padrões previstos para Salvador, um com final baixo, L + H* L%, para a mulher e L + H* H% para o homem. Do ponto de vista fonológico, as duas realizações confirmam o padrão H + L* L% para as declarativas (vermelho) e o padrão L + H* L% para as interrogativas (azul), proposto

por Moraes (2008), sendo que no homem se observa a variante descrita por Lira (2009) e Silva (2011) com final alto também para Fortaleza.

Na variedade do **Rio de Janeiro**, tanto o homem quanto a mulher realizam nas frases declarativas (em vermelho) um padrão alto e ascendente na pretônica do núcleo (sílabas 7), “pássaro”, descendente na sílaba tônica, o padrão vai descendo suavemente nas sílabas postônicas que não são elididas.

Figura 11 – Variedade Rio de Janeiro: escolaridade média

Enunciado 2: O pássaro gosta do pássaro./ O pássaro gosta do pássaro?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Rio de Janeiro

B: F0 do homem de Rio de Janeiro

No pré-núcleo “pássaro”, o comportamento do padrão é bem diferente do comportamento no núcleo, ver sílabas 2, 3 e 4, para o pré-núcleo. Trata-se de um comportamento ascendente na sílaba tônica, sílaba 2, que continua subindo suavemente pelas duas sílabas posteriores, 3 e 4. Não há apagamento de sílabas e as curvas se realizam progressivamente. Este comportamento rítmico e melódico, tão diferente dos encontrados nos falantes de Salvador e Fortaleza, poderia ser atribuído ao nível diferenciado de escolaridade e formação destes dois falantes. Ambos demonstram ter adquirido um padrão de leitura rítmico, bem como padrões melódicos mais regulares para a oralização dos dados de leitura.

Do ponto de vista fonológico as duas realizações confirmam o padrão H + L* L% para as declarativas (vermelho) e o padrão L + H* L% para as interrogativas (azul), proposto por Moraes (2008).

Em termos de duração, nas proparoxítonas se repete o mesmo padrão de supressão silábica para Fortaleza. Entretanto, tanto Salvador como Fortaleza apresentam maior diferença entre as sílabas tônicas e átonas no pré-núcleo “pássaro” – sílabas 2, 3, 4 – e “gosta” – sílabas 5 e 6 – do que o Rio de Janeiro. Estas observações rítmicas qualitativas

precisariam ser sistematicamente descritas posteriormente a partir de uma análise quantitativa.

Figura 12 – Duração de enunciado com núcleo proparoxítono: mulheres

Enunciado 2: O pássaro gosta do pássaro./ O pássaro gosta do pássaro?
Declarativas (vermelho) e interrogativas (azul).
Duração das mulheres de Salvador, Fortaleza, e Rio de Janeiro:

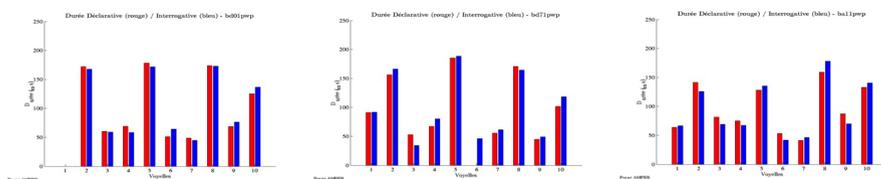
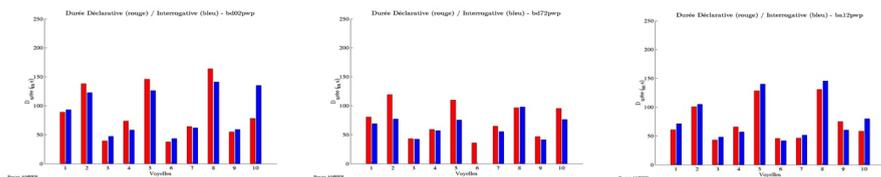


Figura 13 – Duração de enunciado com núcleo proparoxítono: homens

Enunciado 2: O pássaro gosta do pássaro./ O pássaro gosta do pássaro?
Declarativas (vermelho) e interrogativas (azul).
Duração dos homens de Salvador, Fortaleza, e Rio de Janeiro:



Nas proparoxítonas se repete o mesmo padrão de supressão silábica nos falantes homens de Salvador e Fortaleza frente ao padrão do homem do Rio de Janeiro, que apresenta menos diferença de duração entre tônicas e átonas, com um contorno mais proporcional.

Em termos de intensidade, o planejamento da leitura com a queda abrupta de intensidade no final do enunciado também se realiza no padrão proparoxítono da falante do Rio de Janeiro.

Figura 14 – Intensidade de enunciado com núcleo proparoxítono: mulheres

Enunciado 2: O pássaro gosta do pássaro./ O pássaro gosta do pássaro?
 Declarativas (vermelho) e interrogativas (azul).
 Intensidade das mulheres de Salvador, Fortaleza, e Rio de Janeiro:

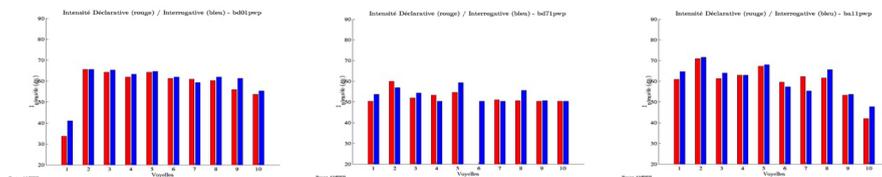
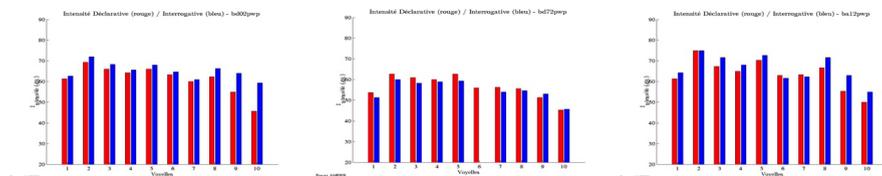


Figura 15 – Intensidade de enunciado com núcleo proparoxítono: homens

Enunciado 2: O pássaro gosta do pássaro./ O pássaro gosta do pássaro?
 Declarativas (vermelho) e interrogativas (azul).
 Duração dos homens de Salvador, Fortaleza, e Rio de Janeiro:



Nas localidades Salvador e Rio de Janeiro, a sílaba tônica das proparoxítonas, sílaba 8 do núcleo, a intensidade da interrogativa (azul) é maior do que a da declarativa (vermelho). Em Fortaleza, a mulher repete o comportamento das outras duas localidades, mas, no homem, a intensidade da declarativa é maior que a da interrogativa.

O planejamento da leitura com a queda abrupta progressiva de intensidade no final do enunciado também se realiza no padrão proparoxítono do falante do Rio de Janeiro.

4.3 Resultados e discussões: núcleos oxítonos

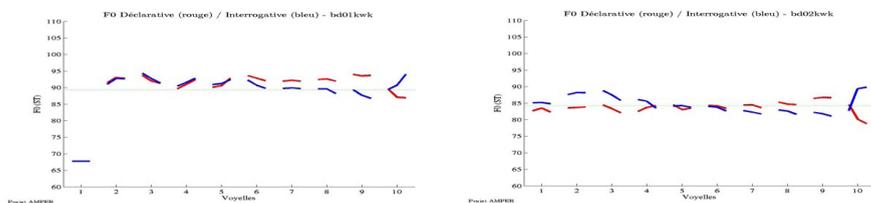
A categoria acentual oxítona é a categoria mais polêmica do português em termos de classificação ou atribuição acentual para o núcleo, que pode ser interpretado como um padrão truncado ou comprimido. O padrão acentual das oxítonas é analisado a partir dos três parâmetros

acústicos: F0, duração e intensidade em enunciados declarativos e interrogativos totais.

Em termos de F0, na variedade de **Salvador**, tanto o homem quanto a mulher realizam nas frases declarativas um tom alto na pretônica do núcleo, “bisavô”, ver sílaba 9 (em vermelho). A última sílaba tônica é descendente, sendo mais baixa que a pretônica. Isso confirma o acento tonal proposto para o núcleo das declarativas por Moraes (2008), sem o tom de fronteira, em H + L* L% de declarativas.

Figura 16 – Variedade de Salvador: escolaridade média

Enunciado 3: O bisavô gosta do bisavô./ O bisavô gosta do bisavô?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Salvador

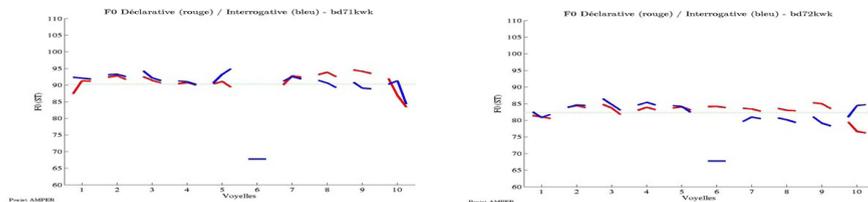
B: F0 do homem de Salvador

No pré-núcleo o comportamento do padrão é bem diferente do comportamento do núcleo, ver sílabas 2, 3 e 4. Trata-se de um comportamento alto ou ascendente progressivo, sem a queda do núcleo. Nas interrogativas o padrão também pode ser considerado como truncado, uma vez que temos uma pretônica baixa, sílaba 9 em azul, seguida de uma tônica alta, sílaba 10. O padrão proposto por Moraes (2008) para o núcleo das interrogativas totais L + H* L% se realiza sem o tom de fronteira baixo, pois na nossa interpretação não há segmento posterior para a realização do tom baixo de fronteira L%, terminando numa sílaba acentuada alta H*, por truncamento.

Na variedade de **Fortaleza**, tanto o homem quanto a mulher realizam nas frases declarativas um tom alto na pretônica do núcleo, “bisavô”, ver sílaba 9 (em vermelho). A última sílaba tônica é descendente, sendo mais baixa que a pretônica. Isso confirma o acento tonal proposto para o núcleo das declarativas por Moraes (2008) H + L* L%.

Figura 17 – Variedade de Fortaleza: escolaridade média

Enunciado 3: O bisavô gosta do bisavô./ O bisavô gosta do bisavô?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Fortaleza

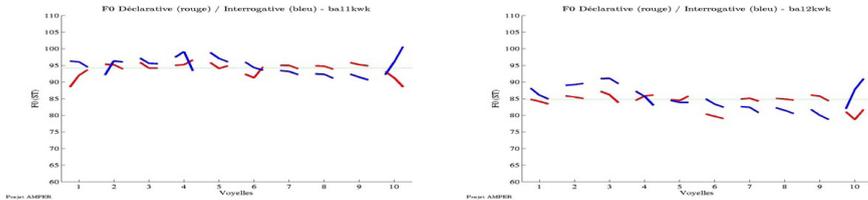
B: F0 do homem de Fortaleza

No pré-núcleo o comportamento do padrão é bem diferente do comportamento do núcleo, ver sílabas 2, 3 e 4. Trata-se de um comportamento alto H*, sem a queda longa e inclinada do núcleo. Nas interrogativas o padrão também pode ser considerado como comprimido, uma vez que temos uma pretônica baixa, sílaba 9 em azul, seguida de uma tônica alta, sílaba 10, ascendente-descendente para a mulher e ascendente com um princípio de descida para o homem. O padrão proposto por Moraes (2008) para o núcleo das interrogativas totais L + H* L% se realiza com o tom de fronteira baixo para a mulher, mas não para o homem que termina numa sílaba alta, conforme o assinalado por Lira (2009, p. 95), para o padrão oxítono de Fortaleza, por truncamento ou compressão.

Na variedade do **Rio de Janeiro**, tanto o homem quanto a mulher realizam nas frases declarativas um tom alto descendente na pretônica do núcleo, “bisavô”, ver sílaba 9 (em vermelho), sendo esta pretônica o pico do enunciado. A última sílaba tônica é descendente, sendo mais baixa que a pretônica. Isso confirma o acento tonal proposto para o núcleo das declarativas por Moraes (2008), sem o tom de fronteira, em H + L* L% implementando-se a descida na tônica quando esta é a última sílaba do enunciado.

Figura 18 – Variedade Rio de Janeiro: escolaridade média

Enunciado 3: O bisavô gosta do bisavô./ O bisavô gosta do bisavô?
Declarativas (vermelho) e interrogativas (azul).



A: F0 da mulher de Rio de Janeiro

B: F0 do homem de Rio de Janeiro

No pré-núcleo o comportamento do padrão é bem diferente do comportamento do núcleo, ver sílabas 2, 3 e 4. Trata-se de um comportamento alto ou ascendente progressivo, sem a queda do núcleo. Nas interrogativas o padrão também pode ser considerado como truncado, uma vez que temos uma pretônica baixa, sílaba 9 em azul, seguida de uma tônica alta, sílaba 10. O padrão proposto por Moraes (2008) para o núcleo das interrogativas totais $L + H^*$ $L\%$ se realiza sem o tom de fronteira baixo, pois na nossa interpretação não há segmento posterior para a realização do tom baixo de fronteira $L\%$, terminando numa sílaba acentuada alta H^* , por truncamento.

Em termos de duração, a duração das vogais tônicas nucleares tende a ser mais longa nas sílabas tônicas das interrogativas (azul) do que nas declarativas (vermelho), pelo menos nos dados de Salvador e Fortaleza, mas não nos do Rio de Janeiro. O mesmo comportamento de duração se encontra nos homens, o falante de Fortaleza é o que mais tende à supressão silábica, sendo que nos três casos a distribuição dos padrões de redução ou alongamento da duração vocálica em dois grupos rítmicos simétricos aponta para uma oralização mais planejada das frases lidas.

Figura 19 – Duração de enunciado com núcleo oxítono: mulheres

Enunciado 3: O bisavô gosta do bisavô./ O bisavô gosta do bisavô?
 Declarativas (vermelho) e interrogativas (azul).
 Duração das mulheres de Salvador, Fortaleza, e Rio de Janeiro

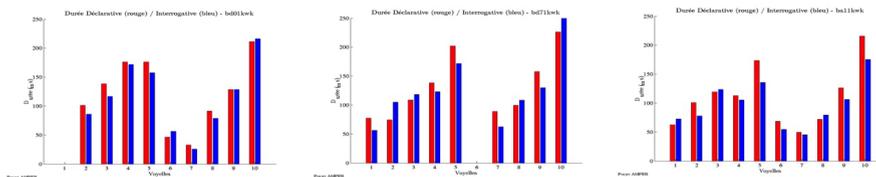
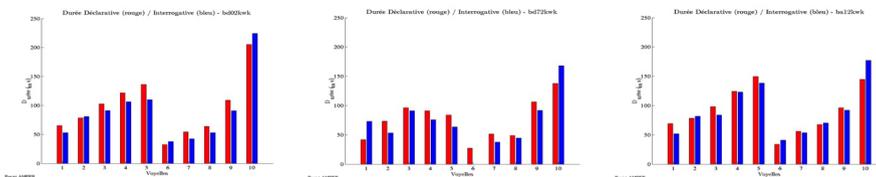


Figura 20 – Duração de enunciado com núcleo oxítono: homens

Enunciado 3: O bisavô gosta do bisavô./ O bisavô gosta do bisavô?
 Declarativas (vermelho) e interrogativas (azul).
 Duração dos homens de Salvador, Fortaleza, e Rio de Janeiro



Nas sentenças com núcleo oxítona, o padrão de supressão recai sobre as sílabas átonas internas do enunciado, “gosta do”, sílabas 6 e 7, também para a mulher de Fortaleza. A falante do Rio de Janeiro que mantém o seu padrão rítmico de leitura, criando mesmo dois grupos rítmicos simétricos no interior do enunciado, fazendo o contraponto entre “gosta” e “bisavô”, sílabas 5 e 10. O mesmo realiza a falante de Fortaleza apesar da supressão silábica em “gosta” e menos a falante de Salvador que apresenta dois grupos rítmicos também de leitura, mas menos regulares ou simétricos que os de Fortaleza e Rio de Janeiro.

Em termos de intensidade, no caso do Rio de Janeiro, a queda abrupta de intensidade se dá mesmo no padrão oxítono, no caso das declarativas (vermelho), estratégia de leitura que se repete no falante masculino, como se verá a seguir.

Figura 21 – Intensidade de enunciado com núcleo oxítono: mulheres

Enunciado 3: O bisavô gosta do bisavô./ O bisavô gosta do bisavô?
Declarativas (vermelho) e interrogativas (azul).
Intensidade das mulheres de Salvador, Fortaleza, e Rio de Janeiro

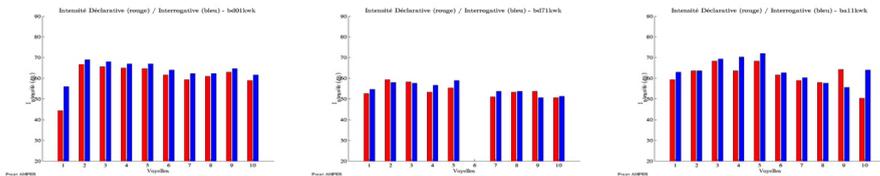
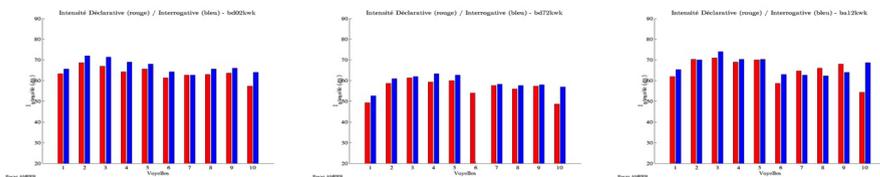


Figura 22 – Intensidade de enunciado com núcleo oxítono: homens

Enunciado 3: O bisavô gosta do bisavô./ O bisavô gosta do bisavô?
Declarativas (vermelho) e interrogativas (azul).
Intensidade dos homens de Salvador, Fortaleza, e Rio de Janeiro



Assim como para as mulheres, no caso do Rio de Janeiro, a queda abrupta de intensidade se dá mesmo no padrão oxítono, nas declarativas (vermelho), estratégia de leitura que associada ao comportamento de duração e de F0 diferencia os falantes do Rio de Janeiro pela aprendizagem de um padrão prosódico de leitura que é aplicado à oralização dos dados, no momento da leitura do corpus.

5 Considerações finais

Procuramos estabelecer um diálogo entre os resultados de nossas análises e dados de projetos de pesquisa similares para as três capitais. Em termos de duração, a distribuição regular da duração ao longo do enunciado, formando grupos rítmicos simétricos, aponta para a realização de uma fala planejada, leitura, no caso das três localidades, o que atribuímos ao grau de escolaridade dos falantes, que já adquiriram a prosódia de leitura, aprendida nas práticas escolares. A internalização

de um padrão mais silábico e regular de leitura dos falantes de Salvador, Fortaleza e do Rio de Janeiro também se destaca no caso das proparoxítonas e das oxítonas.

No caso dos homens, em termos de duração, a distribuição das reduções e alongamentos vocálicos por sílaba também é muito regular, tanto nos dados do Rio de Janeiro quanto nos de Salvador e menos nos de Fortaleza. Os resultados referentes à intensidade não podem ser considerados, sem normalização, em termos brutos, pois as condições de gravação certamente não foram as mesmas, o que afeta o dado em dB. Entretanto, podemos comparar a distribuição dos valores de intensidade ao longo do enunciado, comparando as diferenças e proporções de intensidade interssilábicas, uma vez que são as mesmas frases e que a variação no interior do enunciado da intensidade é desprezível.

Assim como para as mulheres, nos dados dos homens, para as três localidades, a intensidade da sílaba tônica no núcleo é mais importante na interrogativa (azul) do que na declarativa (vermelho). A partir dos dados analisados, encontramos diferenças significativas em relação às modalidades – declarativa e interrogativa total – e, em menor medida, em relação à origem dos falantes.

5.1 Com relação às modalidades podemos afirmar que:

- a. A pretônica que antecede a última tônica é mais alta na modalidade declarativa do que na interrogativa total;
- b. A duração da última tônica é menor na modalidade declarativa que na interrogativa total, menos para a variedade Salvador;
- c. A intensidade da última tônica é menor na modalidade declarativa que na interrogativa total;
- d. O contorno entonativo da modalidade declarativa é descendente, ao passo que o da interrogativa total é, predominantemente, circunflexo proparoxítono e paroxítono.
- e. No padrão oxítono das interrogativas, temos resultados ascendentes para o Rio de Janeiro, truncados para Salvador (ascendente com começo de queda) e circunflexo na mulher, mas truncado no homem de Fortaleza.

5.2 Com relação à origem dos falantes podemos afirmar que:

- a. Nas declarativas, o comportamento da F0 é descendente da pretônica nuclear, descendendo gradualmente, até a postônica (linha de declinação). Entretanto, nas interrogativas, o movimento de F0 é sempre ascendente na sílaba tônica nuclear, nas três localidades. Fortaleza diverge de Rio de Janeiro e Salvador, pois a subida da F0 na tônica é seguida de movimento descendente na própria tônica.
- b. A diferença de duração entre sílabas tônicas e átonas em posição nuclear agrupa Rio de Janeiro e Salvador em oposição ao comportamento de Fortaleza. Em Fortaleza a tônica nuclear é proporcionalmente mais longa do que a postônica que chega inclusive à elisão vocálica na mulher.
- c. A queda de intensidade ao longo do enunciado é visível nas três variedades, no entanto, a queda no núcleo da tônica para a postônica é mais abrupta em Fortaleza devido aos apagamentos finais.
- d. O contorno entonativo da modalidade declarativa é descendente nas três localidades, ao passo que o da interrogativa total é predominantemente circunflexo nos núcleos paroxítonos e proparoxítonos. Nos núcleos oxítonos é variável. Segundo sugestão de Moraes (2016, informação verbal) seria importante verificar se o padrão ascendente das oxítonas é um caso de truncamento ou um padrão diferente do padrão das paroxítonas através de testes perceptuais com estímulos gerados a partir da manipulação de síntese em pontos chave dos enunciados.

Agradecimentos

Agradecemos a João Antônio de Moraes e Zulina Souza de Lira por terem disponibilizado os dados de Salvador e Fortaleza para esta análise. E a Albert Rilliard pelas observações e sugestões teórico-metodológicas para a análise e revisão final do texto. Igualmente, o trabalho atento e cuidadoso, a leitura construtiva dos dois revisores. Quaisquer imprecisões ou erros ainda remanescentes são de responsabilidade nossa, inteiramente.

Referências

- ANTUNES, L. Os projetos AMPER e ALIB e o estudo da variação prosódica no estado de Minas Gerais. In: GSCP INTERNATIONAL CONFERENCE, 2012, Belo Horizonte. *Arquivos...* Belo Horizonte: UFMG, 2012. Disponível em: <http://150.164.100.248/gscp2012-eng/data1/arquivos/leandra_antunes.pdf>. Acesso em: 17 de janeiro de 2017
- ARANTES, P. Time-normalization of Fundamental Frequency Contours: a Hands-On Tutorial. In: MEIRELES, A. (Org.). *Courses on Speech Prosody*. Newcastle upon Tyne: Cambridge Scholars Publishing, 2015. v. 1, p. 98-123.
- BARBOSA, P. A. From syntax to acoustic duration: A dynamical model of speech rhythm production, *Speech Communication*, v. 49, p. 725-742, 2007. <https://doi.org/10.1016/j.specom.2007.04.013>.
- BLANCHE-BENVENISTE, C. Lo hablado y lo escrito. In: ____ *Estudios lingüísticos de la relación entre oralidad y escritura*. Tradução de Lía Varela. Barcelona: Gedisa, p. 29-63, 1998.
- BOERSMA, P. WEENINK, D. (1992-2014) «PRAAT: doing phonetics by computer». Versão 5.3.84. Disponível em: <www.praat.org>. Acesso em: 18 de setembro de 2016.
- CINTRA, G. Distribuição de padrões acentuais no vocábulo em português, *Confluência*, v. 5 n. 3, p. 82-93, 1997.
- CONTINI, M.; ROMANO, A.; MOUTINHO, L. DE C.; FERNANDEZ REI, E. L'Avancement des recherches en Géoprosodie et le projet AMPER. *Estudios de Fonética Experimental, EFE*, v. XVIII, p. 109-122, 2009. Disponível em: <<file:///C:/Users/Notebook/Downloads/XVIII-11.pdf>>. Acesso em: 20 de setembro de 2016.
- CUNHA, C. S. Entoação regional no português do Brasil. 2000. Tese (Doutorado em Língua Portuguesa) –Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2000.
- CUNHA, C. S. Atlas linguístico do Brasil: uma análise das questões de prosódia. In: MOTA, J. A.; CARDOSO, S. M. (Org.). *Documentos 2: Projeto Atlas Linguístico do Brasil*. Salvador: Editora Quarteto, 2005. v.1. p.187-205.

CUNHA, C. S. Corpus ALIB: uma base de dados para pesquisas atuais e futuras. In: CUNHA, C. S. (Org.) *Estudos geo-sociolinguísticos*. Rio de Janeiro: Programa de Pós Graduação em Letras Vernáculas, UFRJ, 2006. p. 67-81.

GOMES DA SILVA, C. *Análise entonacional e pragmática de conversas telefônicas coloquiais: os enunciados interrogativos totais nas variedades de Buenos Aires e Santiago do Chile*. 2014. Dissertação (Mestrado em Letras Neolatinas) – Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014. Disponível em: <<http://www.lettras.ufrj.br/pgneolatinas/media/bancoteses/carolinagomesmestrado.pdf>>. Acesso em: 20 de março de 2017.

LADD, D. Robert. *Intonational Phonology*. 2. ed. Cambridge: Cambridge University Press, 2008. <https://doi.org/10.1017/CBO9780511808814>.

LIRA, Z. de. *A entoação modal em cinco falares do Nordeste brasileiro*. 2009. Tese (Doutorado em Linguística) - UFPB, João Pessoa, 2009. Disponível em: <http://bdtd.biblioteca.ufpb.br/tde_busca/arquivo.php?codArquivo=449>. Acesso em: 20 de março de 2017.

MIRANDA, L. S. *Análise da entoação do português do Brasil segundo o modelo IPO*. 2015. Dissertação (Mestrado em Língua Portuguesa) – Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015. Disponível em: <<http://www.lettras.ufrj.br/posverna/mestrado/MirandaLS.pdf>>. Acesso em: 20 de março de 2017.

MORAES, J. A de; ABRAÇADO, M. J. A. A descrição prosódica do português do Brasil no AMPER, *Geolinguistique – Hors série*, n. 3, p. 337-345, 2005.

MORAES, J. A. de. The pitch accents in Brazilian Portuguese: analysis by synthesis. *Speech Prosody*, v. 4, p. 389-397, 2008. Disponível em: <<http://sprosig.isle.illinois.edu/sp2008/papers/8inv.pdf>>. Acesso em: 15 de dezembro de 2016.

MORAES, J. A. de. *Data Analysis*. Rio de Janeiro: Curso de Pós Graduação em Letras Vernáculas e Neolatinas da UFRJ, 2016.

MOUTINHO, L.; COIMBRA, R. L.; RILLIARD, A.; ROMANO, A. Mesure de la variation prosodique diatopique en portugais européen. *Estudios de fonética experimental*, v. 20, p. 33-55, 2011.

NUNES, V. G. *A prosódia de sentenças interrogativas totais nos falares catarinenses e sergipanos*. 2015. Tese (Doutorado) – Programa de Pós Graduação em Lingüística da Universidade Federal de Santa Catarina, Florianópolis, 2015.

REIS, C. *L'Interaction entre l'Accent, l'Intonation et le Rythme em Portugais brésilien: Étude Acoustique de la Prosodie*. 1995. Tese (Doutorado) – Institut de Phonétique, Université de Provence, Aix-en-Provence, 1995.

RILLIARD, A.; LAI, J.-P. Outils pour le calcul et la comparaison prosodique dans le cadre du projet AMPER – l'exemple des variétés Occitane et Sarde. In: SYMPOSIUM INTERNATIONAL SUR LA VARIATION DIATOPIQUE DE L'INTONATION DANS LE DOMAINE ROUMAIN ET ROMAN, 2008, Iasi, Roumanie. *Actes...* Iasi, Roumanie: Editura Universitatii «Alexandru Ioan Cuza», 2008. p. 217-229.

RILLIARD, A. La base de données AMPER. *Géolinguistique, hors-série*, n. 4, 2011.

RILLIARD, A. *Prosodie et Interaction Homme-Machine: Étude de la variation démarcative, diatopique, diachronique & expressive*. 2014. Mémoire (Habilitation à Diriger des Recherches de l'Université Paris-Sud) - Université Paris-Sud, Paris, 2014.

ROMANO, A.; CONTINI, M.; LAI, J. P.; RILLIARD, A. Distancias prosódicas entre variedades románicas en el marco del proyecto AMPER. *Revista Lingüística Iberoamericana*, v. IX, n. 17, p. 13-26, 2011. Disponível em: <<http://www.iber-america.net/cgi-bin/search.cgi?n=verlagskatalog&lang=en&kat=RILI>>. Acesso em: 15 de dezembro de 2016.

ROSSI, Mario. Peut-on prédire l'organisation prosodique du langage spontané?. *Etudes de linguistique appliquée*, v. 66, p. 20-48, 1987.

SEARA, I; REBOLLO-COUTO, L. Entoação de frases declarativas e interrogativas totais no falar fluminense e catarinense. In: CONGRESO INTERNACIONAL DE LA ASOCIACIÓN DE LINGÜÍSTICA Y FILOLOGÍA DE LA AMÉRICA LATINA (ALFAL), XVI., 2011, Alcalá de Henares (Espanha), *Actas...* Alcalá de Henares: ALFAL, 2011. v. 1, p. 1-7.

SILVA, J. C. B. *Caracterização prosódica dos falares brasileiros: as orações interrogativas totais*. 2011. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011. Disponível em: <<http://www.letas.ufrj.br/posverna/mestrado/SilvaJCB.pdf>>. Acesso em: 20 de março de 2017.

SILVESTRE, A. P. S. *A entoação regional dos enunciados assertivos nos falares das capitais brasileiras*. 2012. Dissertação (Mestrado em Língua Portuguesa) - Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2012. Disponível em: <<http://www.letas.ufrj.br/posverna/mestrado/SilvestreAPS.pdf>>. Acesso em: 20 de março de 2017.

VAISSIÈRE, Jacqueline. Langues, prosodies et syntaxe: Prosodie et syntaxe. *TAL. Traitement Automatique des Langues*, v. 38, n. 1, p. 53-82, 1997.

XU, Y. Transmitting tone and intonation simultaneously – the parallel encoding and target approximation (PENTA) model. In: INTERNATIONAL SYMPOSIUM ON TONAL ASPECTS OF LANGUAGES: WITH EMPHASIS ON TONE LANGUAGES, 2004, Beijing, 2004. Disponível em: <http://www.phon.ucl.ac.uk/home/yi/yispapers/Xu_TAL2004.pdf>. Acesso em: 15 de dezembro de 2016.

Desambiguação de sentenças na interface fonologia-sintaxe: resultados de um estudo de compreensão¹

Sentence disambiguation in the phonology-syntax interface: comprehension study results

Melanie Campilongo Angelo²

Universidade de São Paulo, São Paulo, São Paulo / Brasil
melanie.angelo@usp.br

Raquel Santana Santos³

Universidade de São Paulo, São Paulo, São Paulo / Brasil
raquelss@usp.br

Resumo: Este artigo analisa a compreensão no português brasileiro da pista prosódica de duração de sílabas em sentenças ambíguas do tipo SN1-V-SN2-Atributo (e.g. ‘A mãe encontrou a filha suada’). Fonologicamente, tais leituras podem ser explicadas pelo fato de o atributo poder ou não se juntar a seu núcleo na construção do domínio da frase fonológica (NESPOR; VOGEL, 1986), e de que, se há fronteira de domínios, um alongamento é esperado (FOUGERON; KEATING, 1997). O estudo propõe que o alongamento é um fenômeno opcional no PB. Um experimento de *picture matching* é aplicado para versões de estruturas de oposição não local e local de nove sentenças. Os resultados apontaram

¹ Uma versão anterior deste trabalho foi defendida em Angelo (2016). Agradecemos aos participantes da banca de mestrado e a dois pareceristas anônimos pelos comentários e discussão do texto e os eximimos de todos os problemas remanescentes.

² Agradece o auxílio em Forma de bolsa de Mestrado do Departamento de Linguística da FFLCH/USP (CAPES Proex 2013-2015).

³ Agradece o auxílio do CNPq (Bolsa Produtividade 308135/2009-1).

para diferenças significativas conforme o tipo de estrutura sintática. As diferentes sentenças também se mostraram relevantes para a interpretação de que, ainda que o alongamento seja opcional na produção, uma vez realizado, ele serve como condutor para uma interpretação não local. Os resultados para uma interpretação local parecem decorrer da interação entre o Princípio de *Late Closure* (FRAZIER, 1979) com a estrutura prosódica da sentença. Por fim, a análise evidencia que estruturas do tipo *small clause* interferem no processo de alongamento e, assim, este artigo defende que isso ocorre porque a reestruturação prosódica é bloqueada neste tipo de estrutura sintática.

Palavras-chave: posição local; posição não local; fronteira prosódica; alongamento; *small clauses*.

Abstract: This article discusses the comprehension of syllable duration in Brazilian Portuguese as a prosodic cue in ambiguous sentences with a NP1-V-NP2-attribute structure (e.g. ‘The mother has found her daughter sweating’). Phonologically, interpretations of the above-mentioned structure can be explained by the fact that attributes may or may not join the head in the construction of the phonological phrase domain (NESPOR; VOGEL, 1986), and because lengthening is expected when there is a boundary (FOUGERON; KEATING, 1997). We suggest that lengthening exists in BP as an optional phenomenon. We ran a picture matching experiment, with versions of structures with high and low attachment of 9 sentences. Overall results showed significant differences for type of syntactic structure. The different sentences also played a role in the results, indicating that although lengthening is optional in production, once it has been performed, it leads to a high attachment reading. The Low Attachment Principle (FRAZIER, 1979) seems to play a role in the results for sentences with low attachment. Finally, the findings signal also that sentences which allow small clause constructions may interfere with the lengthening process; and we argue that this happens because restructuring is blocked in this kind of structure.

Keywords: low attachment; high attachment; prosodic boundary; lengthening; small clauses.

Recebido em: 10 de dezembro de 2016.

Aprovado em: 17 de janeiro de 2017.

1 Introdução

Neste artigo, investigamos o uso da pista prosódica da duração de sílabas na desambiguação de sentenças do Português Brasileiro (doravante PB), através de escolhas de interpretação do ouvinte de contextos onde pode haver reestruturação de frases fonológicas. Nosso objetivo é observar se a maior/menor duração de um trecho guia o ouvinte a uma dada interpretação e se, em caso positivo, há gradações no *continuum* duracional. Para tanto, elegemos casos de ambiguidade sintática permanente, especificamente sentenças de tipo SN1-V-SN2-Atributo, tais como (1). Nosso interesse está no papel que a prosódia tem na desambiguação.

Magalhães e Maia (2006) avaliaram a interpretação da leitura de sentenças que apresentam ambiguidade entre as posições local/não-local do atributo, como em (1), que podem ter as leituras (a) com aposição não-local, e (b) com aposição local.

1. O pai abraçou o filho embriagado.
 - a. O pai estava embriagado.
 - b. O filho estava embriagado.

Os autores aplicaram um experimento em 2 tarefas: (a) os informantes liam as sentenças e depois diziam qual a interpretação, e (b) liam as sentenças com algum tipo de informação que pudesse influenciar a estrutura prosódica das sentenças (barra (/) entre ‘filho’ e ‘embriagado’). Os resultados apontaram que, em frases onde não havia nenhuma marca que pudesse influenciar a prosódia, havia sempre uma preferência pela aposição local. Quando havia barra entre o objeto e o atributo, aumentava a quantidade de opções pela interpretação de aposição não-local. Quando a leitura era pela aposição não-local, havia um alongamento na sílaba tônica do atributo (a sílaba ‘ga’).

Fonseca e Magalhães (2008) apontam também outras pistas prosódicas que interferem na escolha da interpretação destas sentenças. No experimento de leitura em voz alta, encontrou-se uma elevação da frequência fundamental em SN1 e no Atributo e uma pausa silenciosa entre SN2 e o Atributo. Fonseca (2008), por sua vez, observou o uso de aspectos entoacionais neste mesmo tipo de estrutura, observando se algum padrão específico é adotado no momento em que o falante

necessita desambiguar uma sentença. Os resultados encontrados foram de que apenas quando a prosódia é enfática (elevação de F0 no sujeito e no atributo) a interpretação se direciona para o não-local.

Angelo e Santos (2015) reformularam os testes aplicados por Magalhães e Maia (2006) e analisaram dados de produção para observar o que acontecia com a duração no trecho onde pode haver reestruturação prosódica – duas frases fonológicas se juntarem na composição de um único domínio prosódico – uma vez que o resultado apresentado pelos autores é interessante por não poder ser explicado pelas teorias fonológicas atuais. Como veremos, na leitura não-local, ‘filho’ e ‘embriagado’ estão em domínios prosódicos diferentes ([o filho ϕ] [embriagado ϕ]), enquanto que em uma leitura local, ‘filho’ e ‘embriagado’ estão em um mesmo domínio prosódico ([o filho embriagado ϕ]).

Angelo e Santos tomam como pressuposto resultados translinguísticos que apontaram que sílabas tônicas e sílabas finais de palavras são alongadas no final de domínios prosódicos (OLLER 1973; KLATT, 1976; KEATING *et al.*, 2003). Se este é o caso, o fato de ‘ga’ se tornar mais longo em Magalhães e Maia não deveria favorecer nenhuma interpretação, pois em ambas as interpretações ‘ga’ é a sílaba tônica final de um domínio de frase fonológica. Além disso, alguns estudos também constataram que quanto mais alto o domínio prosódico, maior é a duração (cf. TABAIN, 2003; KEATING *et al.*, 2003). Neste caso, a previsão seria de uma variação na duração das sílabas de ‘filho’ apenas na aposição não-local, por serem próximas à fronteira de frase fonológica. Ou seja, deveria ocorrer um alongamento nas sílabas do SN2, e não do atributo, pois é ele que se encontra em diferentes domínios prosódicos conforme a interpretação. Um alongamento nas primeiras sílabas do atributo na interpretação não-local, como no caso da sílaba ‘em’ de ‘embriagado’, também se justificaria por estar próximo à fronteira. Já na leitura com aposição local, tais sílabas não estão no final/começo de domínio e, portanto, deveriam ser mais curtas do que com a leitura não-local, quando estão no começo/final do domínio.

Os resultados de Angelo e Santos (2015) não apresentaram diferença de duração significativa entre as leituras, embora tenha sido encontrada uma clara tendência: sempre que houve um alongamento relevante ele ocorria nos casos de aposição não-local (fronteira). A pergunta que se coloca, então, é se esse alongamento seria opcional. Um experimento de compreensão ajuda a definir em que medida uma duração

maior nas sentenças do tipo descrito em (1a) (ainda que não obrigatória na produção) conduziria o ouvinte a interpretá-las como não-local.

Mas ainda outra questão se coloca. Tanto Magalhães e Maia (2006) quanto Angelo e Santos (2015) trataram as sentenças em (1) como possíveis de ter apenas duas estruturas. No entanto, Angelo (2016) aponta que algumas dessas sentenças podem ter uma terceira estrutura sintática, como em (2):

2. A mãe encontrou a filha suada.
 - a. A mãe estava suada.
 - b. A filha estava suada.
 - c. A mãe se deparou com uma situação: a filha suada.

A interpretação de (2c) é gerada por uma terceira estrutura sintática que leva a um mapeamento prosódico diferente daquele das sentenças locais por adjunção. Esta constatação traz importantes consequências para a expectativa de interpretação das sentenças, já que então não é esperada nenhuma diferença fonológica entre os tipos de leitura local.

Assim, o objetivo deste artigo é discutir o que acontece com as sílabas dentro *versus* às margens de domínios prosódicos em tais sentenças, pois, uma vez que há diferentes mapeamentos estruturais (aliado ao fato de haver na literatura trabalhos que concluem que sílabas em início e/ou final de domínios prosódicos são mais longas e melhor articuladas), espera-se que quando a duração das sílabas que beiram a possível fronteira for maior, a interpretação será não-local, evidenciando que há fronteira entre objeto e atributo (pois a reestruturação não é possível). Também esperamos observar se há uma interpretação diferente do significado das sentenças caso haja uma diferença de estrutura sintática nas leituras locais.

Este artigo organiza-se da seguinte maneira: na seção 2, apresentamos um breve resumo das teorias sintáticas, fonológicas e de processamento, no que dizem respeito especificamente ao que se sabe sobre estruturas ambíguas. As seções 3 e 4 trazem o *design* do experimento e os resultados de sua aplicação, respectivamente. A discussão dos resultados é apresentada na seção 5, e as considerações finais na seção 6.

2 Sentenças SN1-Verbo-SN2-Atributo nas interfaces Fonologia-Sintaxe-Processamento

2.1 Sintaxe: Ambiguidade por aposição de atributo e Teoria X-Barra

Do ponto de vista sintático, a ambiguidade das sentença SN1-V-SN2-Atributo diz respeito a quais elementos formam um ou mais constituintes. No entanto, essa ambiguidade pode depender da natureza do verbo em conjunto com o complemento que o segue. Segundo Foltran (1999), quando o predicativo é expresso por adjetivos, ele pode ser um adjunto adnominal. Sendo assim, em uma sentença como (4), em que só há aposição local, o adjetivo funciona como modificador do sintagma nominal [o carro].

4. O João comprou o carro quebrado

Alguns testes de gramaticalidade, como passivação, topicalização e/ou clivagem podem nos dar evidências empíricas para a existência de outras estruturas e, então, de uma nova ambiguidade (cf. (5), extraído de Foltran. 1999, p. 29). Fica evidente que no Grupo 1, ‘quebrado’ é um adjunto de ‘carro’ (integra o constituinte por ele nucleado). No entanto, há uma outra construção da sentença, em que, a partir das mesmas construções, o adjetivo ‘quebrado’ comporta-se como um constituinte à parte do sintagma nominal ‘o carro’ mas ainda o caracterizando, só que de forma predicativa – cf. (6).

5. Grupo 1: Uso atributivo do adjetivo

- a. O carro quebrado foi comprado por João.
- b. O carro quebrado, o João comprou-o.
- c. Foi o carro quebrado que o João comprou.

6. Grupo 2: Forma predicativa do adjetivo

- a. O carro foi comprado quebrado por João.
- b. O carro, o João comprou-o quebrado.
- c. Foi o carro que o João comprou quebrado.

No entanto, nem todos os verbos permitem essas duas possibilidades: (8) e (9) trazem os mesmos testes acima para a sentença (7), apenas para a interpretação local:

7. O pai visitou o filho feliz.

8. Grupo 1: Uso atributivo do adjetivo
 - a. O filho feliz foi visitado pelo pai.
 - b. O filho feliz, o pai visitou-o.
 - c. Foi o filho feliz que o pai visitou.

9. Grupo 2: Forma predicativa do adjetivo
 - d. #O filho foi visitado feliz pelo pai.⁴
 - e. #O filho, o pai visitou-o feliz.
 - f. #Foi o filho que o pai visitou feliz.

Como se observa, tanto (4) quanto (7) permitem a interpretação local; no entanto, elas diferem quanto à estrutura sintática em questão, pois ao passo que a primeira permite leitura predicativa e adjuntiva (cf. (5)-(6)), a segunda não é tão clara quanto à estrutura predicativa (cf. (8)-(9)). O que está em pauta é que há uma diferença entre sentenças como (10) e (11) abaixo

10. A Maria trabalhou magoada.

11. O João considera a Maria bonita.

Em (10), o verbo ‘trabalhar’ só seleciona um argumento, o de sujeito, no caso [A Maria], o que não exclui a relação evidente entre o sintagma adjetivo (AP) [magoada] e o sintagma determinante (DP) [A Maria]. Já no segundo exemplo, o verbo ‘considerar’ seleciona, além do sujeito, o constituinte [a Maria bonita] como um todo, sendo uma evidência para a classificação de [a Maria bonita] como uma pequena

⁴ # indica uma sentença gramatical, mas com um significado diferente do pretendido.

oração selecionada pelo verbo, ou seja, uma *small clause* (SC) – cf. Foltran (1999). Nos dois casos, no entanto, temos uma oposição local.

2.2 Fonologia Prosódica: Fronteiras prosódicas e ambiguidade

A discussão da interação entre fonologia e sintaxe não é nova (e.g. LIGHTFOOT, 1976; CHOMSKY; LASNIK, 1978), embora diferentes propostas se apresentem para explicar como esta interação se dá.

Segundo Selkirk (1984) e Nespor e Vogel (1986), parte do componente fonológico interage com outros componentes gramaticais indiretamente, através da criação dos domínios onde as regras fonológicas se aplicam: regras de formação de domínios levam em conta informações de outros componentes gramaticais. Uma vez criados estes domínios, a fonologia já não tem acesso aos outros componentes. Estes domínios (conhecidos como domínios prosódicos) estão estruturados hierarquicamente como uma árvore de ramificação n-ária. De acordo com Nespor e Vogel, a estrutura sintática é mapeada fonologicamente no nível da frase fonológica (ϕ) por meio das regras de mapeamento em (12):

12. Phonological Phrase formation:

- I. Domain: The domain of ϕ consists of a clitic group (C) which contains a lexical head (X) and all Cs on its nonrecursive side up to the C that contains another head outside of the maximal projection of X.
- II. Construction: Join into an n-ary branching ϕ all Cs included in a string delimited by the definition of the domain of ϕ .

ϕ Restructuring (optional):

A nonbranching ϕ which is the first complement of X on its recursive side is joined into the ϕ that contains X.

(NESPOR; VOGEL, 1986, p.168-173),

Uma palavra lexical juntamente com seus clíticos forma o grupo clítico (C) (e.g. ‘o filho’). Um adjetivo, além de compor ele mesmo uma frase fonológica, pode ser incorporado ao domínio que contém a palavra que ele modifica em um processo de reestruturação: anexar a uma frase fonológica o primeiro complemento de X que esteja em seu lado recursivo,

ou seja, o lado em que normalmente se encontram os complementos do núcleo lexical. Para a reestruturação ocorrer, o complemento deve ser uma frase fonológica formada por apenas um grupo clítico, o que impede que sequências como ‘comeu bolo de chocolate’ sejam reestruturadas, já que o complemento do verbo ‘comer’ é ‘bolo de chocolate’, formado por dois grupos clíticos.

As línguas variam quanto a permitir a reestruturação ou não, e em caso afirmativo, deve-se definir se esta reestruturação é obrigatória ou opcional. O PB vem sendo descrito como uma língua em que é possível a reestruturação (cf. ABOUSALH, 1997; SANTOS, 2003; SÂNDALO; TRUCKENBRODT, 2002), embora não se discuta a questão da obrigatoriedade do processo.

O PB é identificado como língua com recursividade à direita. Assim, o adjetivo é uma frase fonológica independente (porque sua projeção máxima é diferente da projeção máxima do nome), mas que pode se reestruturar compondo uma única frase fonológica. Este mapeamento reflete diferenças estruturais de sentenças ambíguas de adjunção, como em (13), entre leituras locais vs. não-locais. Na interpretação de que o filho é feliz, ‘feliz’ é complemento de ‘filho’ e, portanto, as duas frases fonológicas podem ser reestruturadas (13a); na interpretação de que o pai é feliz, não há relação entre ‘filho’ e ‘feliz’, portanto a reestruturação não é permitida entre [o filho] e [feliz] (13b):⁵

13. O pai visitou o filho feliz.

a. leitura: O filho feliz.

[o pai ϕ] [visitou ϕ] [o filho ϕ] [feliz ϕ]

>> [o pai ϕ] [visitou ϕ] [o filho feliz
 ϕ reestruturado]

b. leitura: O pai feliz.

[o pai ϕ] [visitou ϕ] [o filho ϕ] [feliz ϕ]

>> *[o pai ϕ] [visitou ϕ] [o filho feliz
 ϕ reestruturado]

>> [o pai ϕ] [visitou o filho ϕ reestruturado]
 [feliz ϕ]

⁵ Mas note-se que é possível entre o verbo e seu complemento, criando ‘visitou o filho’.

2.3 Processamento: A compreensão de sentenças ambíguas

Estudos em Psicolinguística propõem que somos dotados de uma espécie de ‘dispositivo’ (*parser*, analisador sintático) que determina a estrutura de uma sentença, contribuindo para a produção e compreensão da linguagem. A Teoria do *Garden Path* (FRAZIER, 1979) é um modelo de processamento de frases que tem como características ser universal, inato, e estar baseado em princípios de economia cognitiva, ou seja, um funcionamento com menor custo e menor memória de trabalho (cf. FRAZIER; FODOR, 1978; FRAZIER, 1979). Segundo este modelo, as escolhas de interpretação são feitas no momento em que se dá o processamento, obedecendo os seguintes princípios:

Minimal Attachment: Attach incoming material into the phrase-marker being constructed using the fewest nodes consistent with the well-formedness rules of the language under analysis. (FRAZIER, 1979, p. 24)

Late Closure: When possible, attach incoming lexical items into the phrase or clause currently being parsed. (FRAZIER, 1979, p. 33)

Conforme o primeiro princípio, ao ouvirmos/lermos uma sentença, o *parser* constrói a estrutura de uma sentença ouvida/lida com o mínimo de nós sintáticos possíveis. Devido ao princípio de *Late Closure*, novos constituintes devem ser apostos ao sintagma que está sendo processado, ou seja, o nó não terminal mais baixo possível. Isto significa que o fechamento do sintagma em processamento é atrasado para permitir que novos itens sejam integrados localmente. Para exemplificar, vejamos (14):

14. Enquanto as meninas costuravam as meias caíram.
- a. Enquanto as meninas costuravam as meias / caíram.
 - b. Enquanto as meninas costuravam / as meias caíram.

O leitor, ao chegar ao sintagma ‘as meias’, interpreta-o como complemento do verbo ‘costuravam’. Isso decorre do princípio de *Late Closure*: o sintagma verbal está aberto para incluir o material que vem a seguir, o DP ‘as meias’ (14a). No entanto, ao encontrar mais material

(‘caíram’), ele reanalisa a sentença encerrando o sintagma verbal (*Early Closure*), de forma que o sintagma ‘as meias’ possa ser analisado como sujeito da próxima oração e não como objeto da anterior (14b).

Uma das direções de trabalho da pesquisa em processamento consiste em descobrir com que tipo de informações o *parser* trabalha. Fodor (1998, 2002a) observa que, nos princípios propostos por Frazier (1979), não havia menção à interferência da prosódia, mas não há como negar que a prosódia tem um papel fundamental no processamento de sentenças, principalmente tendo em vista a quantidade de trabalhos apontando para o seu papel na desambiguação de sentenças (e.g. LEHISTE, 1973). Tendo como base uma diferença nos resultados de resolução de ambiguidades para o inglês e o espanhol em situação de leitura, que pode ser explicada como devida à diferente sensibilidade ao tamanho de constituintes (FODOR, 1998), a autora propõe que a prosódia está presente inclusive em situações de não produção sonora das sentenças, como no caso de leituras silenciosas, e que esta projeção mental da estrutura prosódica (prosódia implícita, nas palavras da autora) é tratada como parte do *input* pelos leitores. Em linhas gerais, esta é a assunção da Hipótese da Prosódia Implícita (Fodor, 2002b), que sugere a existência de uma prosódia que ajuda na resolução de ambiguidades sintáticas na leitura silenciosa, assim como a prosódia explícita faz na fala:

Implicit Prosody Hypothesis: In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction. (FODOR, 2002a, p. 1)

Esta prosódia implícita, assim como a prosódia explícita, tem as características prosódicas de cada língua específica. De acordo com a proposta da HPI, características prosódicas particulares das línguas são responsáveis pelas variações encontradas nas preferências por posições do *parser* – como por exemplo a acima mencionada diferente sensibilidade ao tamanho de constituinte (FODOR, 1998).

De acordo com Miyamoto (1999), o português brasileiro é uma língua com preferência de processamento para posição local. O autor testou a preferência de anexação de orações relativas reduzidas

ou não, anexadas em posição local (baixa) ou não-local (alta), como exemplificado em (15) (exemplo (7) de Miyamoto, 1999):

15. a) A Kombi trouxe os supervisores do engenheiro [que foram pagos pela empreiteira.
- b) A Kombi trouxe o supervisor dos engenheiros [que foram pagos pela empreiteira.
- c) A Kombi trouxe os supervisores do engenheiro [pagos pela empreiteira.
- d) A Kombi trouxe o supervisor dos engenheiros [pagos pela empreiteira.

Os resultados apontaram, no entanto, para uma interação significativa entre o tipo de oração relativa e o local de aposição. A performance com orações relativas reduzidas era melhor com aposição local do que não-local. No caso de orações relativas plenas, a performance foi melhor com a aposição local apenas numericamente, sem significância estatística.

Esta tendência à aposição baixa/local é interpretada como uma tendência a um *Late Closure* (em oposição a línguas com tendência ao *Early Closure*, como o francês, holandês, espanhol – cf. Fodor, 2002). Miyamoto (2005) reanalisa seus dados de (1999) e argumenta que as diferenças no número (singular e plural dos SN) afetou os resultados.

Ribeiro (2001) encontrou uma preferência por aposição alta, não-local, em experimentos que mediam o tempo de leitura em trechos críticos de sentenças com sintagmas adjetivos. Outros estudos atestaram a relação entre prosódia e aposição. Finger & Zimmer (2005) mostraram que orações relativas longas são mais sujeitas à aposição não-local, alta, que orações relativas curtas. Lourenço-Gomes e Moraes (2005) também encontraram para orações relativas longas uma preferência para aposição não-local. Finalmente, Maia *et al.* (2007), a partir de meta-análise, argumentam que a diferença na preferência entre aposição local e não-local só se dá claramente em testes *off-line*.

O que os estudos acima têm em comum é que a aposição não-local é preferida sempre em condições específicas, como a oração relativa ser longa. Assim, mantém-se a afirmação inicial de Miyamoto (1999) de que o PB tem preferência por aposição local. De acordo com Magalhães

e Maia (2006), em PB, na ausência de pistas prosódicas há um *default* aplicado pelos falantes. Este *default* seria a aposição local, justificada pelo Princípio do *Late Closure* (FRAZIER, 1979), que sugere que um sintagma só se feche depois de conferir se há um elemento que possa ser apostado a ele em seguida. Isso significa que, quando um ouvinte escuta (13), ele identifica que ‘feliz’ pode ser apostado à ‘filho’ (sintagma ainda em aberto no *parsing* durante o processamento para verificar a existência de adjuntos a sua frente), se fechando logo em seguida.

3 O Experimento⁶

O objetivo deste artigo é analisar se e como, fonologicamente, pistas de duração na desambiguação de sentenças do tipo SN1-Verbo-SN2-Atributo influenciam a escolha de interpretação do ouvinte. Uma vez que estatisticamente os resultados de Angelo e Santos (2015) não confirmaram a hipótese de alongamento das autoras onde há fronteira de frase fonológica, mas sabendo que sempre que houve uma maior duração ela foi em favor da interpretação não-local, as questões que se colocam são por que isso acontece; por que, quando há alongamento, ele ocorre em direção do esperado pelas autoras.

Nossa hipótese é que os falantes utilizam o alongamento para desambiguar sentenças ambíguas sintaticamente. A duração serviria, então, como pista para a interpretação de tais sentenças. Como vimos, na ausência de pistas prosódicas, o princípio de *Late Closure* favorece leituras locais. O alongamento desambiguaria, então, sentenças quando há interpretação não-local. No entanto, resultados de pesquisas anteriores apontam que o alongamento é processo opcional, não-obrigatório (dado inclusive que o falante pode utilizar outros processos para apontar para a interpretação pretendida). Em linhas gerais, o que podemos esperar é que, sempre que o alongamento ocorrer, deve ser em sentenças não-locais e, em experimentos de compreensão, a sentença deve ser interpretada como não-local. Nossa hipótese nos leva às seguintes predições específicas:

⁶ Aprovação do Comitê de Ética para Pesquisas com Seres Humanos deferida pelo Instituto de Psicologia da Universidade de São Paulo e registrada através do CAAE: 45791815.5.0000.5561.

- i. Para **sentenças longas**, dado os resultados de Angelo e Santos (2015), espera-se que os falantes as **interpretem como não-local (A)**.
- ii. Para **sentenças curtas**, sendo o alongamento opcional no PB, haveria uma **variação na interpretação** (entre aposição local e não-local); no entanto, pela HPI, deve-se encontrar uma preferência pela leitura local.

Ademais, uma vez que não há trabalhos prosódicos até o momento que considerem possíveis diferenças estruturais em interpretações de aposição local, assumimos que, em princípio, estruturas adjuntivas e que permitem *small clauses* tenham comportamentos similares (uma vez que processualmente possuem a mesma interpretação (local)). No entanto, sentenças que podem ter interpretação local de *small clause* serão investigadas em grupos separados para que haja um controle. Esperamos, assim, que o alongamento seja ‘acertado’ (interpretado como não-local) sempre que realizado; já nas versões curtas, independentemente de a estrutura ser de adjunção ou de *small clause*, pode haver variações entre acertos e erros (pois o alongamento sendo opcional, respostas não-locais também são possíveis).

3.1 Método

3.1.1. Participantes

Participaram do experimento 30 ouvintes adultos, de nível universitário, nascidos em São Paulo, com idade entre 18 e 50 anos.⁷ Para a leitura das sentenças a serem ouvidas, 50 falantes com o mesmo perfil sociolinguístico previamente gravaram as sentenças – lidas em histórias que conduziam ao significado pretendido.

⁷ Todos os participantes completaram o experimento. Posteriormente, os resultados por informante confirmaram que nenhum dos participantes apresentou comportamento desviante do grupo.

3.1.2. Materiais

As sentenças analisadas neste experimento são as mesmas de Angelo e Santos (2015), lidas em histórias que conduziam às interpretações em questão – cf. Quadro 1 e exemplo de história em (16).⁸

Quadro 1 – Sentenças Analisadas

S1. O pai visitou o filho feliz.
S2. A babá ninou a menina chorando.
S3. O aluno consultou o monitor cismado.
S4. O sobrinho cumprimentou o tio sonolento.
S5. O assessor auxiliou o presidente furioso.
S6. O repórter entrevistou o político sozinho.
S7. A mãe procurou a filha magoada.
S8. A mãe encontrou a filha suada.
S9. O réu encontrou o advogado nervoso.

16. A mãe encontrou a filha suada.

- a) Mãe e filha iriam se encontrar para almoçar no shopping antes da filha viajar. No meio do caminho, o pneu do carro da mãe furou e, como não havia ninguém para ajudá-la, a mãe trocou o pneu sozinha, o que a deixou muito cansada. Ao chegar ao shopping, correu muito para não se atrasar ainda mais. A mãe encontrou a filha suada. No entanto, conseguiu chegar antes que a menina fosse embora.
- b) Sábado era aniversário da avó de Lúcia. Sua mãe avisou para que Lucia não se atrasasse e enfatizou que era uma festa chique e Lucia deveria ir bem vestida. Mas naquele dia, a garota tinha vários compromissos: ela saiu cedo de casa, foi trabalhar, depois foi à academia, malhou muito, correu para o shopping para

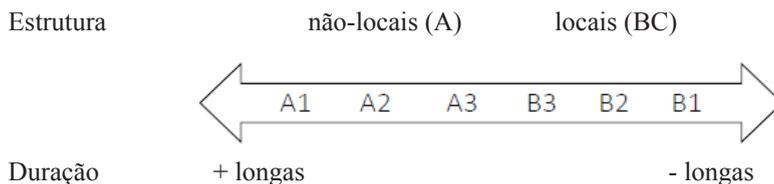
⁸ Antes da aplicação do experimento, as sentenças foram apresentadas para julgamento por 9 estudantes universitários, que não apontaram nenhum problema pragmático ou de impossibilidade de leitura com interpretação local e não-local.

comprar um presente para a avó e teve que ir pra festa sem passar antes em casa pra se arrumar. A mãe encontrou a filha suada. Lucia recebeu uma grande bronca por não estar arrumada direito.

As sentenças apareciam sempre ao final das histórias, mas não necessariamente na última sentença. Todas as sentenças permitiam interpretações não-locais (A) e locais. No caso das sentenças S1 a S7, apenas a interpretação local de adjunção era possível (aqui identificadas como leitura B), ao passo que as sentenças S8 e S9 permitiam, além de adjunção, estrutura de *small clause* (a possibilidade de duas estruturas para a interpretação local é identificada como C). O desbalanceamento na quantidade de sentenças com interpretação B ou C se deve ao fato de que esta diferenciação só ocorreu após a aplicação do experimento.

Para analisar o efeito da duração na interpretação das sentenças era relevante a duração trecho entre a sílaba final do objeto até a sílaba inicial do atributo.⁹ Foram então extraídas para o experimento de compreensão as 3 versões mais longas para o trecho relevante lidas em contexto de interpretação não-local, assim como as 3 versões mais curtas no mesmo trecho lidas em contexto de aposição local, totalizando 54 dados a serem aplicados aos ouvintes.¹⁰ Em uma escala, a gradação das durações se dá conforme representado na Figura 1 abaixo.

Figura 1 – Escala de Gradações das estruturas



⁹ A sílaba tônica não foi alvo de nossa análise porque, como se pode observar nas sentenças, a posição da sílaba tônica variou em relação à fronteira prosódica (em alguns casos estava na fronteira, em outros estava em uma sílaba de distância, e em um caso a duas sílabas de distância) e esta variação afeta os resultados, como já reportado na literatura (cf. Seção 2.2).

¹⁰ Não há um valor mínimo ou máximo de duração especificado para a seleção das sentenças de A1 a B1 pois as sentenças apresentavam segmentos com características acústicas diferentes e não é objetivo desta análise investigar o quanto a duração deve ser mais longa para que se possa atribuir uma dada interpretação.

Os áudios selecionados correspondem à leitura de 17 indivíduos (homens e mulheres) dentre o total de 50 falantes. Os participantes ouvintes tiveram acesso às sentenças, sem nenhum contexto que apontasse para qualquer interpretação.

O material é composto ainda por imagens que foram desenhadas para cada versão das sentenças a serem ouvidas. Em cada imagem, buscou-se evidenciar que somente um personagem não dispunha do atributo que o outro dispunha. Por exemplo, em (7), ora o pai estava feliz e o filho triste, ora o filho estava feliz e o pai triste.

Ao final, o *corpus* é composto por 1620 dados (9 sentenças x 2 interpretações x 3 gradações = 54 respostas de cada um dos 30 ouvintes).

3.1.3 Procedimentos

Antes do teste, os ouvintes fizeram um treinamento elaborado com o programa *PowerPoint*, no qual testava-se se os ouvintes reconheciam os personagens nas imagens e se distinguiam entre as versões não-local e local conforme os atributos das sentenças ('feliz', 'suada', etc). Como o treinamento tratava de identificar a diferença entre os pares de imagens, a noção de ambiguidade fica implícita; por isso, não se utilizou distratoras no teste. O treinamento durava 5 minutos.

Após o treinamento, o teste de *picture matching* era iniciado, elaborado com o software *TP Versão 3.1* da *Worken*.¹¹ Cada *input* do teste correspondia a um dos 54 áudios, ou seja, todos os participantes ouviram as duas versões/interpretações da mesma sentença, e cada uma em três versões de duração, mas de forma aleatorizada, diminuindo, assim efeitos de explicitação/reforço. Na tela, um par de imagens aparecia no lado esquerdo e um botão de ok ao centro do lado direito (portanto, equidistante das duas imagens), em seguida uma sentença era tocada. As imagens foram apresentadas sempre na mesma ordem (versão A acima e versão B/C abaixo), independentemente da resposta esperada para cada áudio. O ouvinte deveria clicar em cima da imagem (resposta) interpretada (versão A ou versão B/C). Para ouvir o áudio seguinte, o ouvinte deveria clicar em 'ok'. Além disso, os ouvintes usaram um fone

¹¹ Agradecemos à prof. Dra. Andreia Rauber, uma das criadoras do TP, pela cessão do programa, disponível gratuitamente através da URL <http://www.worken.com.br/tp_regfree.php>.

confortável e de alta precisão conectado ao computador para fazer o teste. O teste durava, em média, 10 minutos.

Os resultados foram automaticamente codificados pelo programa no *Excel* com as respostas de cada ouvinte identificadas entre acerto ou erro, definidas a depender da intenção do falante ao ler cada sentença no experimento de produção (intenção definida pelo contexto em que sentença se inseria).

4. Resultados

Os resultados do experimento de compreensão foram analisados por tipo de estrutura, por sentença e por informante, no entanto a análise por ouvinte não trouxe inferências significantes ao escopo de nossas hipóteses. A aplicação dos testes estatísticos foi feita através do programa *R*. Para alguns casos, utilizou-se o teste de proporção, em outros, o teste para igualdade de proporções. Reportamos abaixo os resultados por estrutura e sentença.

4.1 Por estrutura

A Tabela 1 abaixo apresenta os valores gerais por tipo de estrutura, comparando as proporções de acertos, sem considerar uma possível influência da sentença. O teste rodado foi o de igualdade de proporções, que busca responder se a diferença das proporções de acertos das estruturas é igual a 0. A comparação de A com B/C tem base em todas as 9 sentenças do experimento, já em A com B, apenas as sentenças de 1 a 7 foram consideradas. Quando se compara A com C, apenas as sentenças 8 e 9 são as observadas.

TABELA 1
 Comparação da proporção de acertos conforme a estrutura

Teste para igualdade de proporções		
Estrutura	Diferença entre as proporções de acerto (IC95%)	p-valor
<i>A</i> <i>B/C</i>	0,02 (-0,03; 0,07)	0,472
<i>A</i> <i>B</i>	0,16 (0,10; 0,21)	<0,001
<i>A</i> <i>C</i>	-0,46 (-0,55; -0,37)	<0,001

Conforme se observa acima, posto que o valor da diferença entre as proporções de acerto subtrai as locais das versões não-locais em cada linha, na comparação geral de A com B/C não se observou diferença significativa na frequência de acertos. No entanto, há diferença na proporção de acertos de A comparada a B – a frequência de acerto é maior para A. Ao comparar A e C, também existe diferença significativa, mas note que, neste caso, a frequência de acertos é maior para C.¹²

A Tabela 2 abaixo registra a quantidade de acertos e erros segundo a estrutura e a gradação de longas e curtas. Como dissemos, em uma escala, A1 se trata da sentença com o trecho relevante mais longo das sentenças não-locais, enquanto A3 se trata da menos longa destas. Ao mesmo tempo, B1 e C1 são as que apresentaram os trechos relevantes mais curtos das sentenças locais, ao passo que B3 e C3 são as menos curtas. Isto significa que as 1 são aquelas que mais têm chances de conduzir o ouvinte ao acerto em ambos os casos (segundo nossas predições), e as 3 são as mais suscetíveis a erro.

¹² Valores de IC positivos indicam maior taxa de acerto para A. Valores negativos apontam para maior taxa de acerto para B ou C.

TABELA 2
 Frequência de acertos e erros conforme a estrutura e a duração
 (gradação de longa e curta)

Teste de Proporções					
Estrutura	Certo		Errado		p-valor
	N	%	N	%	
<i>A</i>	514	63,46	296	36,54	<0,001
<i>A1</i>	196	72,59	74	27,41	<0,001
<i>A2</i>	166	61,48	104	38,52	<0,001
<i>A3</i>	152	56,30	118	43,70	0,045
<i>B</i>	351	55,71	279	44,29	0,004
<i>B1</i>	110	52,38	100	47,62	0,535
<i>B2</i>	124	59,05	86	40,95	0,011
<i>B3</i>	117	55,71	93	44,29	0,113
<i>C</i>	148	82,22	32	17,78	<0,001
<i>C1</i>	43	71,67	17	28,33	0,001
<i>C2</i>	54	90,00	6	10,00	<0,001
<i>C3</i>	51	85,00	9	15,00	<0,001
<i>B/C</i>	499	61,60	311	38,40	<0,001
<i>B1/C1</i>	153	56,67	117	43,33	0,033
<i>B2/C2</i>	178	65,93	92	34,07	<0,001
<i>B3/C3</i>	168	62,22	102	37,78	<0,001

Conforme se pode observar, o número de acertos das estruturas gerais (A, B, C e B/C) foram estatisticamente diferentes de 0,5.¹³ Para A, C

¹³ Um p-valor significativo (<0,05) nos testes que comparam se a diferença entre acertos e erros são diferentes de 0,5 indica que o ouvinte acertou ou errou as respostas, fugindo da média de oscilação. Em paralelo à observação do p-valor, quando este for significativo o leitor deve considerar as porcentagens de acertos e erros, pois mostrarão a direção da significância.

e B/C, obteve-se p -valor $< 0,001$, porém, para a estrutura B encontramos p -valor = 0,004, o que nos mostra que essa é a estrutura menos acertada de forma geral, ainda que com mais acertos do que erros. Analisando a quantidade de acertos, podemos observar que a estrutura C foi a mais acertada – na grande maioria das vezes, ao ouvir uma sentença curta, o ouvinte a interpretava como local. A estrutura B foi a menos acertada (55,7%). Tal resultado indicia a opcionalidade do alongamento desse tipo de estrutura. Sentenças com interpretação local podem ser reestruturadas prosodicamente. Se o são, SN2 não está em fronteira de frase fonológica. Se não há reestruturação, há uma fronteira fonológica logo depois de SN2, como ocorre com as sentenças em interpretação não-local. Finalmente, a estrutura A teve uma taxa de acerto de 63%, o que aponta a um direcionamento para a interpretação não-local, de acordo com o predito: contextos mais longos levariam à interpretação não-local. Mas interessantemente, não é uma previsão nossa que o ouvinte pudesse ouvir sentenças com duração mais longa e interpretá-las como local, o que ocorreu em 36,5% dos casos.

Segundo nossas previsões, as sentenças 1 deveriam apresentar mais acertos. Vejamos os resultados no que concerne às gradações para cada estrutura. As estruturas não-locais (A1, A2 e A3) obedecem a uma ordenação de acertos, pois, embora todas sejam significativas, as mais longas (A1 e A2) apresentam valor de significância menor que A3. Veja que esta gradação é percebida mesmo na percentagem de acertos: A1 teve 72% de acerto, A2 teve 61% e A3 teve 56% de acertos.

Nas gradações das sentenças locais que não permitem a interpretação de *small clause* (B1, B2 e B3), essa ordenação não é encontrada. B2 apresentou mais acertos que B1 e B3. No caso das sentenças que permitiam *small clause*, todas as gradações de C (C1, C2 e C3) apresentaram acertos significativos, embora percentualmente C2 tenha sido a mais acertada – mas ressaltamos que havia apenas 2 sentenças com esse tipo de estrutura. Nas gradações das sentenças locais como um todo, ou seja, considerando tanto aquelas que permitem como as que não permitem *small clauses* (B/C1, B/C2 e B/C3), o resultado foi um p -valor significativo para todas, mas com uma direção oposta ao esperado: B/C1 apresenta um p -valor maior do que B/C2 e B/C3.

4.2 Por sentença

A Tabela 3 abaixo indica a proporção de acertos para cada sentença, relacionando-as aos tipos de estrutura.

TABELA 3

Comparação da proporção de acertos segundo estrutura, para cada sentença

Teste para igualdade de proporções			
Sentença	Estrutura	Diferença entre as proporções de acerto (IC95%)	p-valor
S1	A	0,17	0,035
	B	(0,01; 0,32)	
S2	A	0,09	0,272
	B	(-0,06; 0,24)	
S3	A	0,17	0,025
	B	(0,02; 0,31)	
S4	A	-0,10	0,175
	B	(-0,24; 0,04)	
S5	A	-0,11	0,143
	B	(-0,26; 0,03)	
S6	A	0,43	<0,001
	B	(0,29; 0,58)	
S7	A	0,44	<0,001
	B	(0,31; 0,58)	
S8	A	-0,54	<0,001
	C	(-0,68; -0,41)	
S9	A	-0,38	<0,001
	C	(-0,52; -0,24)	

Como se observa, o fator 'sentença' merece investigação, já que algumas sentenças tiveram as leituras mais claramente identificadas do que outras. Observou-se diferença significativa na comparação entre A e B nas sentenças S1, S3, S6 e S7, sendo a frequência de acertos sempre maior para A (não-local) do que para B (local). As sentenças S2, S4 e S5 não apresentaram diferenças significativas – e no caso de S4 e S5, houve mais acertos para as estruturas locais (B). Quando se compara A e C (sentenças S8 e S9), a diferença também é significativa, mas aqui a frequência de acertos é maior nas estruturas de C.

Lembramos que a ausência de significância na comparação de A vs. B/C para cada sentença é justificada uma vez que esperamos que os ouvintes oscilem entre acertos e erros para as respostas de B ou C. Ou seja, não necessariamente deve haver diferença estatísticas entre os erros/acertos das não-locais vs. locais.

A Tabela 4 considera, para cada sentença, as diferentes estruturas e suas gradações.

TABELA 4
Frequência de acertos e erros conforme a sentença
para cada estrutura e gradações

Teste de proporções						
Sentença	Estrutura	Certo		Errado		p-valor
		n	%	n	%	
<i>S1</i>	A	59	65,56	31	34,44	0,004
	<i>A1</i>	23	76,67	7	23,33	0,006
	<i>A2</i>	16	53,33	14	46,67	0,856
	<i>A3</i>	20	66,67	10	33,33	0,100
	B	44	48,89	46	51,11	0,961
	<i>B1</i>	10	33,33	20	66,67	0,100
	<i>B2</i>	19	63,33	11	36,67	0,201
	<i>B3</i>	15	50,00	15	50,00	1,000

Teste de proporções						
Sentença	Estrutura	Certo		Errado		p-valor
		n	%	n	%	
S2	A	63	70,00	27	30,00	<0,001
	A1	24	80,00	6	20,00	0,002
	A2	27	90,00	3	10,00	<0,001
	A3	12	40,00	18	60,00	0,361
	B	55	61,11	35	38,89	0,045
	B1	15	50,00	15	50,00	1,000
	B2	20	66,67	10	33,33	0,100
S3	A	69	76,67	21	23,33	<0,001
	A1	24	80,00	6	20,00	0,002
	A2	25	83,33	5	16,67	<0,001
	A3	20	66,67	10	33,33	0,100
	B	54	60,00	36	40,00	0,073
	B1	17	56,67	13	43,33	0,584
	B2	18	60,00	12	40,00	0,361
S4	A	61	68,89	28	31,11	<0,001
	A1	21	70,00	9	30,00	0,045
	A2	22	73,33	8	26,67	0,018
	A3	19	63,33	11	36,67	0,201
	B	71	78,89	19	21,11	<0,001
	B1	27	90,00	3	10,00	<0,001
	B2	22	73,33	8	26,67	0,018
B3	22	73,33	8	26,67	0,018	

Teste de proporções						
Sentença	Estrutura	Certo		Errado		p-valor
		n	%	n	%	
S5	A	58	64,44	32	35,56	0,008
	A1	23	76,67	7	23,33	0,006
	A2	15	50,00	15	50,00	1,000
	A3	20	66,67	10	33,33	0,100
	B	68	75,56	22	24,44	<0,001
	B1	22	73,33	8	26,67	0,018
	B2	23	76,67	7	23,33	0,006
	B3	23	76,67	7	23,33	0,006
S6	A	62	68,89	28	31,11	<0,001
	A1	25	83,33	5	16,67	<0,001
	A2	22	73,33	8	26,67	0,018
	A3	15	50,00	15	50,00	1,000
	B	23	25,56	67	74,44	<0,001
	B1	9	30,00	21	70,00	0,045
	B2	6	20,00	24	80,00	0,002
	B3	8	26,67	22	73,33	0,018
S7	A	76	84,44	14	15,56	<0,001
	A1	26	86,67	4	13,33	<0,001
	A2	26	86,67	4	13,33	<0,001
	A3	24	80,00	6	20,00	0,002
	B	36	40,00	54	60,00	0,073
	B1	10	33,33	20	66,67	0,100
	B2	16	53,33	14	46,67	0,855
	B3	10	33,33	20	66,67	0,100

Teste de proporções						
Sentença	Estrutura	Certo		Errado		p-valor
		n	%	n	%	
S8	A	23	25,56	67	74,44	<0,001
	A1	12	40,00	18	60,00	0,361
	A2	3	10,00	27	90,00	<0,001
	A3	8	26,67	22	73,33	0,018
	C	72	80,00	18	20,00	<0,001
	C1	18	60,00	12	40,00	0,361
	C2	27	90,00	3	10,00	<0,001
	C3	27	90,00	3	10,00	<0,001
S9	A	42	46,67	48	53,33	0,598
	A1	18	60,00	12	40,00	0,361
	A2	10	33,33	20	66,67	0,100
	A3	14	46,67	16	53,33	0,855
	C	76	84,44	14	15,56	<0,001
	C1	25	83,33	5	16,67	<0,001
	C2	27	90,00	3	10,00	<0,001
	C3	24	80,00	6	20,00	0,002

Quando se estratifica a frequência de acertos por estrutura para cada sentença, observa-se que com exceção das sentenças S8 e S9 todas apresentam valores significativos e com mais acertos que erros nas estruturas de A. É interessante notar que essa significância se encontra, na maioria das vezes nas sentenças A1 e A2 (que são as mais longas). Não houve nenhum caso em S1 a S7 de que uma sentença menos longa (A3) tenha sido significativa e suas versões mais longas não. No caso das sentenças S8 e S9, o comportamento foi diferente: S8 apresentou A significativo (que só ocorreu em A2), e com maior número de erros do que de acertos, enquanto que para S9 nenhum resultado foi significativo nas gradações de A.

Os resultados encontrados para as gradações das versões B não foram tão claros. S1, S3 e S7 não apresentam diferenças significativas em nenhuma das gradações. No caso de S2, apenas a frequência geral de acertos da gradação foi significativa. As sentenças S4 e S5 apresentam valores significativos em todas as gradações. A sentença 6 apresenta um padrão diferente já que todas as gradações foram significativas, porém, os falantes erraram mais em todos estes os casos.

Nas gradações das sentenças C, que permitem *small clause* (S8 e S9), com exceção de C1 na S8, todos os valores foram significativos. Do mesmo modo que nas sentenças não-locais, não houve nenhum caso em que as gradações 2 ou 3 tenham sido significativas e as respectivas gradações mais curtas (1 ou 2) não.

Considerando apenas os valores das estruturas gerais, nas estruturas de A (não-local longas), com exceção das sentenças S8 e S9 (que permitem interpretação de *small clause*), todas apresentam p-valor significativo com acertos em direção ao esperado (não-local). A sentença S8 apresenta p-valor significativo, mas para números de erros (ou seja, mesmo quando essa sentença é longa, os ouvintes preferem a interpretação local). Já a sentença S9 não apresenta p-valor significativo para A, indicando que os ouvintes oscilam nas escolhas de respostas para essa sentença.

A respeito da estrutura geral B, as sentenças S1, S3 e S7 não apresentaram significância estatística. Nas sentenças S2, S4 e S5, assim como nas sentenças S8 e S9, que permitem *small clause*, o p-valor é significativo para o número de acertos. A sentença S6, diferentemente, apresenta significância em B, mas para o número de erros, ou seja, mesmo quando a sentença é curta, os falantes preferem a interpretação não-local.

5 Discussão

Nossa primeira predição, dado os resultados de Angelo e Santos (2015), era de que as sentenças em que o trecho analisado foi alongado pelos falantes na produção seriam interpretadas como aposição não-local (A).

Considerando os acertos e erros conforme a estrutura, vimos que os acertos de A foram significativamente maiores que os acertos de B. Ainda, é interessante notar que esse resultado se manteve não só na estrutura geral de A, mas em todas as suas gradações, sendo A3 (mais curta entre as longas) com menor p-valor que A1 e A2.

Ao mesmo tempo, para os resultados por sentenças, à exceção das sentenças S8 e S9, todas apresentaram acertos significativos nas versões A: ou seja, quando os ouvintes escutaram as sentenças com alongamento, de fato preferiram a versão A. Estes resultados também vão ao encontro das propostas de que há alongamento quanto maior for o domínio prosódico (CHO; KEATING, 2001; KEATING *et al.*, 2003).¹⁴

A segunda predição foi levantada esperando que o processo de alongamento exista como opcional no PB. Assim, haveria uma variação na interpretação (entre aposição não-local e local) das sentenças em que não houvesse alongamento: curtas seriam igualmente interpretadas como local e não-local.

Cabe analisar, primeiramente, os acertos e erros por estrutura (sem contrapô-las às interpretações opostas). No geral, todas as estruturas foram mais acertadas que erradas. Como vimos na Tabela 2, os acertos foram significativos em todas as gradações de duração de A, assim como nas gradações de C. Interessantemente, em B só houve significância no valor geral e na segunda gradação, mas vale ressaltar que os p-valores não foram tão baixos como em A e C. Assim, as gradações em que não houve significância de acertos em B confirmam a predição (que previa que os acertos seriam iguais a 50%, ou seja, oscilação de respostas). Quando unimos B/C há significância, mas pode ser pela influência dos altos índices de significância de C.

Pela interação com a HPI, esperávamos que, em caso de oscilação, os ouvintes optassem pela interpretação local. No entanto, os resultados de B não confirmam esta predição, pois quando não houve alongamento, os ouvintes nem sempre responderam em direção à interpretação local (como deveria ser caso o alongamento fosse obrigatório), em B1 e B3 há variação nas respostas; no geral (B), o valor é significativo, mas menor que em A.

Olhando para as sentenças, encontramos dois comportamentos. As estruturas curtas das sentenças S1, S3 e S7 não apresentaram significância. As respostas para as versões curtas (B) ficaram na média de 50% de acertos e erros. Conforme a Tabela 3, vimos que há muito mais oscilações nas respostas de B do que nas A, evidenciando que o alongamento parece ser opcional e, por isso, há variações de respostas

¹⁴ Perceptualmente vão contra os achados de Santos e Leal (2008), que não encontraram alongamento no PB em sentenças sem ambiguidade, em um experimento de produção.

nas versões curtas ouvidas, corroborando a sugestão de opcionalidade do processo de Angelo e Santos (2015).

Porém, outras sentenças apresentaram significância de acertos (S2, S4, S5, S8 e S9) da interpretação B, corroborando Magalhães e Maia (2006), pois na ausência de pistas prosódicas, mais especificamente quando não houve alongamento, além das oscilações de respostas, observamos uma preferência pela interpretação de aposição local em 4 das 9 sentenças.¹⁵ Fodor (2002) propõe que esses princípios podem não ser cumpridos na leitura se houver pistas prosódicas da língua que direcionem à outra interpretação. Uma vez que a autora compara a prosódia implícita aos padrões de prosódia explícita, mesmo nosso experimento não se tratando de leitura (mas audição), podemos inferir que a ausência de alongamento é uma ausência de pista prosódica explícita, o que direciona o falante ao *default* local, corroborando Frazier (1979) e Magalhães e Maia (2006).

Para a interpretação B, fica por ser explicada S6 (p-valor < 0,001, mas com relação ao número de erros), que mesmo quando curta, teve preferência por interpretação não-local.¹⁶ Este resultado vai contra as expectativas de uma preferência por aposição local na ausência de pistas prosódicas, nos termos de Magalhães e Maia (2006), mas não contra a proposta fonológica aqui apresentada, pois a interpretação não-local neste caso é possível já que o alongamento é opcional. Porém, esse valor de ‘erro’ foi significativo, diferentemente dos outros casos em que houve variação ou preferência pelas locais. Embora nenhum ouvinte tenha reportado problema de interpretação ou pragmático com essa sentença, nos perguntamos se o ouvinte relacionou o AP ‘sozinho’ ao DP ‘o repórter’, pois seria mais provável pragmaticamente que o repórter estivesse sozinho. Alternativamente, é possível que, por ser a única sentença em que havia 3 personagens nos desenhos, o ouvinte ficasse mais confuso.

¹⁵ Um dos pareceristas apontou que S2, S4, S5 e S8 podem ser enviesadas em termos de conhecimento de mundo. Embora o parecerista explicitasse sua leitura apenas para S2, parece que ele entende que a interpretação nestes casos, por razões pragmáticas, é sempre local. Quando observamos os resultados das gradações na Tabela 4, no entanto, as únicas sentenças com clara direção para a aposição local são S5 e S8. No caso de S8, há a explicação alternativa, desenvolvida na seção 5.1.

¹⁶ Já as versões longas de S6 se comportaram como o esperado, ou seja, também foram interpretadas como não-local.

As sentenças com C, ambíguas entre leitura de adjunto ou *small clause*, apresentaram preferência significativa pela aposição local, podendo indicar apenas a regularidade do princípio de Aposição Local. No entanto, devemos nos perguntar se a estrutura sintática poderia ser uma explicação alternativa para os resultados.

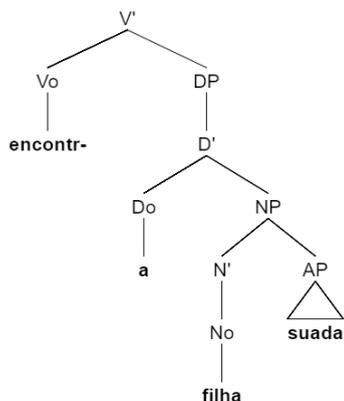
5.1 As *small clauses*, a fonologia prosódica e sentenças ambíguas

Como vimos, as sentenças que permitiam apenas interpretação local por adjunto apresentaram um comportamento diferente daquelas que foram identificadas como ambíguas quanto a interpretação local ser por adjunto ou *small clause*. S1 a S7 (A) foram significativamente interpretadas como não-local. S8 (A) foram significativamente preferidas com interpretações locais; em S9 houve oscilação nas respostas.

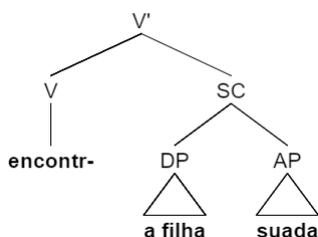
Interessa-nos que as duas estruturas sintáticas locais são diferentes. Segundo Foltran e Miotto (2007), em uma estrutura de adjunção o adjetivo está dentro de um DP, onde é adjunto de um sintagma nominal (NP) – cf. (17a). No caso das *small clauses*, o adjetivo é um predicado (predicativo) de um argumento. Se o argumento é um DP, o adjetivo não pertence a ele, mas forma com ele uma *small clause* – cf. (17b).¹⁷

17. A mãe encontrou a filha suada.

a. Adjunção: [_{DP} a filha suada]



b. *Small clause*: [_{SC} a filha suada]



¹⁷ Para uma proposta de análise distinta das *small clauses*, cf. Starke (1995).

Em (17a) ‘suada’ é complemento de ‘filha’ (ambos derivam de um NP). Em (17b), o adjetivo ‘suada’ não pende de um NP, mas de uma outra categoria (SC). Neste caso, o DP é o sujeito e o AP é o predicado da SC (*small clause*).

Tendo estas duas estruturas em mente, voltemos ao algoritmo de reestruturação de Nespors e Vogel: os domínios de frase fonológica consistem de um grupo clítico que contém um núcleo lexical X e todos os outros grupos clíticos no seu lado não recursivo (à esquerda em PB) até o próximo grupo clítico fora da projeção máxima de X.

A proposta de Nespors e Vogel é de 1986, quando não havia distinção na Teoria X-Barra entre complemento e adjunto e, da mesma forma, NP era a projeção lexical máxima ao invés de DP.¹⁸ Hoje, há trabalhos que mostram que adjuntos funcionam como complementos (e.g. SANTOS, 2003) e diversos trabalhos reveem Nespors e Vogel (GUIMARÃES, 1997; FROTA, 2000; VIGÁRIO, 2003) à luz dos avanços em sintaxe.

Como se nota abaixo, há um diferente mapeamento nos casos de aposição local. Tomemos a sentença (17), que pode ter interpretação por adjunção ou predicativa. No caso de aposição local por atributo, a reestruturação ocorre, pois, ‘suada’ é complemento de ‘filha’ (cf. (18a)). Porém, no caso das *small clauses*, a reestruturação não pode ocorrer, pois embora a interpretação seja local, em sua estrutura sintática o adjetivo não está inserido na projeção máxima do nome. Não estando dentro da projeção máxima do nome (NP), ele não pode ser reestruturado com este – cf. (18b).¹⁹ Por esta organização, há uma fronteira prosódica entre ‘filha’ e ‘suada’, o que nos leva a esperar que a duração seja a mesma que nos casos de aposição não-local (19):

¹⁸ Cf. Chomsky (1970, 1986) para a Teoria X-Barra em vigor na época.

¹⁹ Acrescentamos que a reestruturação é bloqueada tanto se forem uma projeção lexical (STOWELL, 1983) quanto funcional (STARKE, 1995). Se a SC for uma projeção lexical, sua estrutura é similar à de adjunção, porém, ainda assim, ‘suada’ não é complemento de ‘filha’ e isto é o que bloqueia a reestruturação. Por outro lado, se SC for uma projeção funcional, a reestruturação também não ocorre, pois, a SC comporia um outro domínio (já que a frase fonológica de Nespors e Vogel acontece no domínio NP).

18. aposição local

- a. adjunção: [a mãe_φ] [encontrou_φ] [a filha suada_{φreestruturada}]
 b. *small clause*: [a mãe_φ] [encontrou_φ] [a filha_φ] [suada_φ]

19. aposição não-local: [a mãe] [encontrou a filha_{φreestruturada}] [suada_φ]

No caso de aposição não-local, ‘a filha’ pode ser reestruturado com o verbo, já que é o primeiro complemento do verbo (e formado por um único C). No entanto, a estrutura de *small clause* impede esta reestruturação, já que o complemento do verbo é o SC ‘a filha suada’: embora ‘suada’ não seja adjunto de filha, o SC que é complemento do verbo é formado por dois sintagmas (cf. a estrutura em (18b)). Salientamos, no entanto, que esta diferença não traz consequências para o contexto que estamos analisando, já que podemos resumir as fronteiras dos domínios no contexto relevante da seguinte maneira: (i) aposição local: ‘filha’ em fronteira de frase fonológica; (ii) aposição local por adjunção: ‘filha’ dentro de frase fonológica; (iii) aposição local por SC: ‘filha’ em fronteira de frase fonológica.²⁰

Na discussão das predições, não tratamos das sentenças com possibilidade de leitura por *small clause* e, para analisar estes resultados, lembramos que sentenças que permitem *small clause* podem permitir também a interpretação de adjunção, ou seja, não há como controlar a escolha sintática no processamento do falante em nosso experimento e, por isso, essas sentenças foram olhadas com atenção especial nesta análise.

Voltemo-nos então para a discussão do alongamento, lembrando que estamos mais levantando hipóteses a serem aprofundadas, já que havia apenas 2 sentenças deste tipo em nosso corpus. Se apenas a questão da aposição local estiver em jogo (e estruturas predicativas forem

²⁰ Note-se que há diferença entre aposição não-local e aposição local por SC no contexto verbo-nome 2: em aposição não-local, o verbo encontra-se dentro de uma frase fonológica reestruturada com o N2; em aposição local por adjunção, o verbo está em fronteira de frase fonológica (já que o N2 adjunziu-se ao adjetivo e então não pode se juntar ao verbo); e em aposição local por SC, o verbo está em fronteira de frase fonológica (já que não pode se reestruturar com o N2 porque este está dentro da *small clause* com o predicativo (e para o complemento se reestruturar com o núcleo ele deve ser formado por um único C).

interpretadas da mesma forma que estruturas de adjunção), o esperado é que nas versões curtas continue havendo oscilações ou preferência local em S8 e S9, pois, como o processo de alongamento é opcional, a interpretação das curtas pode corresponder a qualquer estrutura (não-local, *small clause* ou adjunção). Por outro lado, se levarmos em conta apenas o algoritmo de construção de frases fonológicas de Nespor e Vogel (1986), espera-se que as versões longas das sentenças que permitem *small clause* (S8 e S9), mesmo produzidas em contexto de interpretação não-local, tenham respostas que direcionem também para a interpretação local. Levando em conta apenas os domínios prosódicos, não há porque se esperar uma tendência para nenhuma das respostas – acerto (o que significaria seleção da estrutura não-local) ou erro (o que significaria seleção da estrutura local de *small clause* ou adjunção).

Ainda mais interessante se torna a expectativa de acertos em se assumindo que o Princípio de Aposição Local (FRAZIER, 1979) interage com a estruturação prosódica. Neste caso, temos que a expectativa de que como *small clause* e não-local permitem alongamento, e *small clause* é uma aposição local, sentenças mais longas deveriam tender a uma preferência por interpretação local.²¹

Na comparação por estrutura, os acertos de C foram significativamente maiores. Isso nos leva a duas interpretações: Ou os acertos de A para as sentenças *small clause* foram poucos, ou os acertos de C foram muitos. Na primeira possibilidade, pode ser que ouvintes tenham escutado as longas e não acertaram por ser possível o alongamento com interpretação local (a de *small clause*). Na segunda possibilidade, as sentenças curtas direcionariam à interpretação local em grande parte das vezes, o que indicaria que, estranhamente os ouvintes não acessam a estrutura de *small clause*, ou que, interessantemente, as leituras de adjunção desses casos são obrigatoriamente curtas, diferentemente das demais estruturas de apenas adjunção (B) – nas quais verificamos que pode haver variação. Neste caso, talvez isso ocorra justamente para diferenciar interpretação local de *small clause* (que deveria ser longa) da de adjunção (sempre curta). Em outras palavras, para as versões curtas,

²¹ O caso ideal de análise seria aquele em que as sentenças pudessem ser apenas de aposição não-local vs. sentenças em que somente a estrutura de *small clause* fosse permitida. Infelizmente, não é este o caso para as nossas sentenças, mas fica aqui a sugestão para trabalhos futuros.

é possível que a preferência local tenha ocorrido como em algumas sentenças B, porém; como em ambos os casos que permitem *small clause* o p-valor foi $< 0,001$ para o acerto (ouviram curta e preferiram local), nos perguntamos se seria o caso de, por haver mais de uma estrutura para interpretação local, o processo de alongamento ser obrigatório para diferenciá-las, ao passo que não o é na interpretação não-local uma vez que o falante também pode lançar mão de outras estratégicas, como ênfase.

Sobre as versões longas dessas sentenças, porém, a sentença S8 apresentou p-valor significativo para números de erros, ou seja, mesmo quando essa sentença é longa, os ouvintes preferiram a interpretação local. Já a sentença S9 não apresentou p-valor significativo para A, sendo o número de erros e acertos bem próximos a 50%. Na análise por estruturas, questionou-se se os acertos em C foram significativamente maiores que A porque houve muitos erros em A ou porque as C foram mais acertadas. De fato, olhando apenas para as estruturas de S8 e S9, vemos que houve uma distorção do padrão nas versões longas dessas sentenças. Em S8, os ouvintes preferiram C independentemente da duração. Em S9, eles preferiram C ao ouvir sentenças curtas, mas variaram nas respostas ao ouvir sentenças longas. Se mesmo ouvindo sentenças longas, os falantes optaram pelas versões curtas, isto é forte indício de que há uma interpretação local com alongamento, devido à fronteira de frase fonológica entre o objeto e o atributo – já que a estrutura por *small clause* bloqueia a reestruturação.

Ao que tudo indica, o bloqueio da reestruturação em C faz com que as sentenças S8 e S9 longas sejam interpretadas também como local, o que não pode acontecer (e não aconteceu) com as sentenças que permitem apenas estrutura de adjunção. No entanto, duas sentenças são pouco para nos permitir fazer afirmações mais contundentes e sugerimos que um estudo futuro analise apenas sentenças desse tipo em contextos em que haja certeza da estrutura lida/interpretada pelo falante/ouvinte.

6 Considerações Finais

O objetivo deste artigo era fornecer uma reanálise do processo de alongamento em fronteira de frase fonológica em contexto de desambiguação de sentenças do tipo SN1-V-SN2-Atributo, buscando trazer mais luzes sobre a questão da interação entre os componentes

gramaticais. Assumindo que essa interação se dá indiretamente, há na fonologia um componente interpretativo que mapeia informações de outros componentes (no caso, a Sintaxe) em níveis e domínios fonológicos (SELKIRK, 1984; NESPOR; VOGEL, 1986).

Tendo em conta resultados translinguísticos a respeito da realização fonética dos segmentos em fronteiras prosódicas em diferentes línguas, em especial do alongamento na produção de segmentos em fronteiras de domínios prosódicos a depender do nível destes domínios, investigamos como o alongamento poderia ser utilizado para distinguir entre interpretações de sentenças estruturalmente ambíguas, buscando evidências de que o alongamento, quando realizado, seria devido a uma fronteira prosódica existente, percebida pelo ouvinte.

Por meio da aplicação de um experimento de *picture matching*, buscamos interpretar as escolhas de significado de sentenças ambíguas levando em conta também gradações no alongamento. Os resultados, de forma geral, corroboraram as predições levantadas. Mais do que uma tendência, encontramos que o processo de alongamento acontece em PB em contexto de desambiguação de sentenças do tipo SN1-Verbo-SN2-Atributo: As estruturas não-locais apresentaram números significativos de acertos e, por sentenças, com exceção das que permitiam *small clauses*, as versões longas foram significativamente interpretadas como de aposição não-local.

Observamos que o processo nem sempre é necessário para que haja interpretações não-locais (ou seja, às vezes, mesmo ouvindo sentenças curtas, os ouvintes selecionavam a resposta não-local), assegurando, assim, que este se trata de um processo opcional na língua, mas favorecido em situações de necessidade de desambiguação de sentenças.

No caso das interpretações locais, os resultados apontaram para a significância de acertos, embora não tão marcantes quanto para as interpretações não-locais. Além disso, não foi encontrada significância estatística para o acerto na análise das diferentes gradações de duração (B1, B2, B3). Olhando para as sentenças separadamente, parte dos dados foram significativos ao acerto, corroborando o Princípio de Aposição Local proposto para o PB em conformidade com o princípio de *Late Closure*, de Frazier (1979) (cf. MAGALHÃES; MAIA 2006), segundo o qual uma preferência pelo *default* local ocorre na ausência de pistas prosódicas.

Interessantemente, duas sentenças foram identificadas pelos informantes como podendo ter 3 interpretações, e por isso foram codificadas separadamente. Trata-se de sentenças em que a aposição local pode-se dar por adjunção ou por *small clauses*. Por serem ambas locais, a expectativa é que tivessem um mesmo comportamento. No entanto, não foi o que ocorreu. As versões longas foram interpretadas como locais. Nossa proposta é de que, embora com interpretação local, o mapeamento prosódico (decorrente da estrutura sintática) tem papel preponderante nas interpretações. Especificamente, estruturas de *small clause* não podem ter o atributo reestruturado ao N2. Assim, estando em frases fonológicas diferentes, elas apresentam a mesma estrutura prosódica que sentenças de interpretação não-local: os falantes identificavam o alongamento e tinham duas possibilidades de escolha – a interpretação não-local ou a local por *small clause*. Nestes casos, os ouvintes selecionavam a interpretação local por HPI. Em outras palavras, não se trata apenas de uma preferência por aposição local, mas por uma interação entre o mapeamento prosódico e Princípio de Aposição Local.

Estes resultados são interessantes, mas como a terceira estrutura só foi identificada *post-hoc*, não houve um balanço entre a quantidade de estruturas com aposição local (*small clause* vs. adjunção) e seus resultados devem ser tomados como preliminares e indicativos de que mais estudos devam ser conduzidos, controlando-se não só aposição, mas tipo de estrutura sintática.

Referências

ABOUSALH, E. F. *Resolução de choques de acento no português brasileiro*. 157 f. Dissertação (Mestrado em Linguística) – Instituto de Estudos Linguísticos, Universidade de Campinas, Campinas, SP, 1997.

ANGELO, M. C. *Produção e percepção na desambiguação de sentenças sintaticamente ambíguas do português brasileiro através da pista prosódica de duração*. 2016. 215 f. Dissertação (Mestrado) – Universidade São Paulo, São Paulo, 2016.

ANGELO, M. C.; SANTOS, R. S. A prosódia em sentenças sintaticamente ambíguas do Português Brasileiro: Pistas de duração. *Alfa: Revista de Linguística*, UNESP, v. 59, p. 375-403, 2015. <https://doi.org/10.1590/1981-5794-1504-7>.

CHO, T.; KEATING, P. Articulatory strengthening at the onset of prosodic domains in Korean. *Journal of Phonetics*, Los Angeles, CA, v.28, p.155-190, 2001. <https://doi.org/10.1006/jpho.2001.0131>.

CHOMSKY, Noam. Remarks on nominalization. In: JACOBS, R. A.; ROSENBAUM, P. S. (Ed.). *Readings in English transformational grammar*, Boston: Ginn, 1970. p. 184-221.

CHOMSKY, Noam. *Knowledge of language: its nature, origin and use*. New York: Praeger, 1986.

CHOMSKY, N.; LASNIK, H. A Remark on contraction. *Linguistic Inquiry*, Camberra, v. 9, n. 2, p. 268-274, 1978.

FINGER, I.; ZIMMER, M. A preferência de interpretação de orações relativas curtas e longas no português brasileiro. In: MAIA, M.; FINGER, I. (Org.). *Processamento da Linguagem: Série Investigações em Psicolinguística GT de Psicolinguística da ANPOLL*. Pelotas: Educat, p. 111-129, 2005.

FODOR, J. D. Learning to parse? *Journal of Psycholinguistic Research*, Springer Link, n. 27, p. 285-319, 1998.

FODOR, J. D. Prosodic disambiguation in silent reading. *Proceedings of North East Linguistic Society*, University of Massachusetts, Amherst, MA, n. 32, p.113-132, 2002a.

FODOR, J. D. Psycholinguistics cannot escape prosody. In: SPEECH PROSODY CONFERENCE, 2002, Aix-en-Provence, France. Aix-en-Provence: Laboratoire Parole et Langage, 2002b.

FOLTRAN, M. J. *As construções de predicação secundária no português do Brasil: aspectos sintáticos e semânticos*. 1999. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 1999.

FOLTRAN, M. J.; MIOTO, C. A favor das *small clauses* revistas. *Cadernos de Estudos Linguísticos*, Campinas, Unicamp, v. 49, n. 1, 2007.

FONSECA, A. A. *Pistas prosódicas e o processamento de sentenças ambíguas do tipo “SN1-V-SN2-Atributo” do Português Brasileiro*. 2008. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

FONSECA, A. A.; MAGALHÃES, J. O. *A influência de pistas prosódicas no processamento de frases: um estudo da estrutura SN1-V-SN2-ATRIBUTO no português brasileiro*. *Alfa – Revista de Linguística*, São Paulo, v. 33, n. 55, p. 1-18, 2008.

FOUGERON, C.; KEATING, P. Articulatory strengthening at edges of prosodic domain. *Journal of the Acoustical Society of America*, Los Angeles, v. 106, p. 3728-3740, 1997. <https://doi.org/10.1121/1.418332>.

FRAZIER, L. *On comprehending sentences: syntactic parsing strategies*. 1979. Tese (Doutorado) – University of Connecticut [reproduzida por Indiana University Linguistics Club], 1979.

FRAZIER, L.; FODOR, J. The sausage machine: a new two-stage parsing model. *Cognition*, Elsevier, v.6, p. 291-326, 1978. [https://doi.org/10.1016/0010-0277\(78\)90002-1](https://doi.org/10.1016/0010-0277(78)90002-1).

FROTA, S. *Prosody and focusing in European Portuguese*. Phonological phrasing and intonation. New York: Garland Publishing. 2000.

GUIMARÃES, M. Unifying LCA & prosodic phrasing in the minimalist program. In: WORKSHOP ON THEORY OF GRAMMAR: PROBLEMS AT PF AND LF INTERFACE LEVELS, 1997. Campinas: Universidade Estadual de Campinas, 1997. Trabalho apresentado.

KEATING, P.; CHO, T.; FOUGERON, C.; HSU, C. Domain-initial articulatory strengthening in four languages. In: LOCAL, J.; OGDEN, R.; TEMPLE, R. (Ed.). *Phonetic interpretation* (Papers in Laboratory Phonology 6). Cambridge: Cambridge University Press, 2003. p.143-161.

KLATT, D. Linguistics Uses of Segmental Duration in English: Acoustic and Perceptual Evidence. *Journal of Acoustic Society of America*, Los Angeles, v. 59, p.1208-1221, 1976. <https://doi.org/10.1121/1.380986>.

LEHISTE, I. Phonetic disambiguation of syntactic ambiguity. *Glossa*, v. 7, p. 107-122, 1973. <https://doi.org/10.1121/1.1982702>.

LIGHTFOOT, D. Trace Theory and Twice-moved NPs. *Linguistic Inquiry*, Camberra, v. 7, n.1, p. 559-582, 1976.

LOURENÇO-GOMES; M. C.; MORAES, J. Efeitos da duração da vogal tônica na percepção de fronteira prosódica em orações relativas estruturalmente ambíguas. ENCONTRO INTERNACIONAL DA ASSOCIAÇÃO BRASILEIRA DE LINGÜÍSTICA, IV., Brasília. Anais... Brasília: Universidade de Brasília. 2005. p. 88. CD-ROM.

MAGALHÃES, J. O.; MAIA, M. Pistas prosódicas implícitas na resolução de ambiguidades sintáticas: um caso de adjunção de atributos. *Revista da ABRALIN*, Florianópolis, v.5, n.1-2, p.143-167, 2006.

MAIA, M. A. R.; FERNÁNDEZ, E.; COSTA, A.; LOURENÇO-GOMES, M. C. Early and late preferences in relative clause attachment in Portuguese and Spanish. *Journal of Portuguese Linguistics*, v.5/6, p. 227-250, 2007.

MIYAMOTO, E. T. *Relative clause processing in Brazilian Portuguese and Japanese*. 1999. Tese (Doutorado) - MIT, Massachusetts, 1999.

MIYAMOTO, E. T. Orações relativas ambíguas e a homogeneidade do processamento de sentenças. In: MAIA, M.; FINGER, I. (Org.). *Processamento da Linguagem*. Pelotas: EDUCAT, 2005. 535p.

NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris Publications, 1986.

OLLER, K. The effect of position in utterance on speech segment duration in English. *Journal of Acoustic Society of America*, Los Angeles, v. 54, p. 1235-1247, 1973. <https://doi.org/10.1121/1.1914393>.

RIBEIRO, A. J. *Um caso de não aplicação preferencial do princípio de Late Closure*. In: CONGRESSO DAASSEL, IX., Rio de Janeiro, 2001. Artigo lido.

SÂNDALO, F.; TRUCKENBRODT, H. Some notes on phonological phrasing in Brazilian Portuguese. *MIT Working Papers in Linguistics*, Cambridge, MA. v. 42, p. 285-310, 2002.

SANTOS, R. S. Traces, pro and stress shift in Brazilian Portuguese. *Journal of Portuguese Linguistics*, Lisboa, v. 2, n. 2, p. 101-113, 2003.

SANTOS, R. S.; LEAL, E. G. Os domínios prosódicos e a duração de sílaba no português brasileiro. *Revista da ABRALIN*, Florianópolis, v. 5, n. 1, p. 143-167, 2008.

SELKIRK, E. *Phonology and Syntax: the relation between sound and structure*. Cambridge: MIT Press, 1984.

STARKE, M. On the format of small clauses. In: CARDINALETTI, A.; GUAISTI, M. T. (Org.). *Syntax and Semantics: small clauses*. San Diego, California: Academic Press, 1995. v. 28, p. 237-269.

STOWELL, T. Subjects across categories. *The Linguistic Review*, Walter De Gruyter, v. 2, p. 285-312, 1983.

TABAIN, M. Effects of prosodic boundary on /aC/ sequences: articulatory results. *Journal of Acoustic Society of America*, Los Angeles, v. 113, p. 2834-2849, 2003. <https://doi.org/10.1121/1.1564013>.

VIGÁRIO, M. *The prosodic word in European Portuguese*. Berlin: Mouton de Gruyter, 2003. <https://doi.org/10.1515/9783110900927>.

Representation and Processing of the Inflected Infinitive in Brazilian Portuguese: an eye-tracking study

Representação e processamento dos infinitivos flexionados em português brasileiro: um estudo com rastreamento ocular

Marcello Modesto

Universidade de São Paulo, São Paulo, São Paulo / Brasil

modesto@usp.br

Marcus Maia

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil

maia@ufrj.br

Abstract: In this paper, we examine the syntactic representation and processing characteristics of null subjects of inflected nonfinite clauses in Brazilian Portuguese (BP). After reviewing some literature on generative-syntax Control and discussing the peculiarities of BP diachrony, we present an eye-tracking experiment which proves that a control interpretation of null subjects of inflected nonfinite clauses is not only psychologically real in BP, but it is actually the preferred option in a task in which a strict comparison with arbitrary PRO is entertained. We then discuss the implications of the experiment to syntactic theory and the analysis of Control and speculate on the role of third factor explanations in the architecture of human language.

Keywords: generative syntax; sentence processing; nonfinite control; inflected infinitives; Brazilian Portuguese; eye-tracking.

Resumo: Este estudo examina a representação sintática e as características de processamento de sujeitos nulos de orações infinitivas flexionadas em português brasileiro (PB). Após revisão de parte da literatura atual sobre a Teoria do Controle e discussão das peculiaridades da diacronia do PB, o artigo apresenta um experimento envolvendo rastreamento ocular

que comprova que a interpretação controlada do sujeito da infinitiva flexionada é não apenas real psicologicamente, mas também a opção preferida numa tarefa em que se compara a leitura controlada à leitura arbitrária. Discute-se, então, as implicações desse experimento com relação à teoria sintática e à Análise do Controle e especula-se sobre o papel de explicações com base no conceito de terceiro fator na arquitetura da linguagem humana.

Palavras-chave: sintaxe gerativa; processamento de frases; controle não finito; infinitivos flexionados; português brasileiro; rastreamento ocular.

Recebido em: 7 de dezembro de 2017.

Aprovado em: 17 de fevereiro de 2017.

1 Introduction: Brazilian Portuguese and nonfinite inflection

The purpose of this study is to investigate the syntactic representation and the psycholinguistic processing of different readings (controlled, arbitrary and referential) of null subjects of inflected nonfinite complement clauses in Brazilian Portuguese (BP). Portuguese (both the European and the Brazilian varieties) is known for having nonfinite inflection (cf. SILVA NETTO, 1950; MAURER, 1968; QUICOLI, 1996; LEMLE, 1984; NEGRÃO, 1986; RAPOSO, 1987, 1989; LIGHTFOOT, 1991; AMBAR, 1994, 1998; MADEIRA, 1994, SITARIDOU, 2002; MILLER, 2002; SCIDA, 2004; NUNES, 2008; MODESTO, 2011, 2016a). However, nonfinite inflection has remained an under-researched characteristic of the language (especially w.r.t. the Brazilian variety) until Modesto (2010) discussed partial control with nonfinite inflection in BP. Modesto's article fed a lot of research on (partial) control and restructuring (LANDAU, 2013, 2015; GRANO, 2012, 2015; WURMBRAND, 2015; MODESTO, 2016a, 2016b) and motivated a return to the study of nonfinite inflection in EP (cf. SHEEHAN, 2014, 2016; GONÇALVES; SANTOS; DUARTE, 2014). In fact, such latter works attested the judgment-data used in Modesto (2010), i.e. the use of nonfinite inflection giving rise to partial control interpretations, in spoken EP and substandard EP. Yet, Modesto's work has been received by some Brazilian linguists with skepticism.

The views advocated and experimentally confirmed in this work (i.e. that nonfinite inflection is used in BP and that its normal interpretation in complement clauses is a controlled one) contrast sharply with the views expressed in Rodrigues and Hornstein (2013). In fact, Modesto (2010) was a reply to Boeckx and Hornstein (2006), which used data from Rodrigues (2004) to show that a theory of control involving syntactic movement (as in HORNSTEIN, 1999) could not explain several facts about BP syntax. However, as argued by Modesto (2010) (and later in 2016a), Rodrigues' data was both problematic and incomplete, because it did not include data on nonfinite inflection. When taken into consideration, nonfinite inflection presents compelling evidence against the hypothesis that the Movement Theory of Control (MTC) could properly explain all the uses of null subjects of finite and nonfinite clauses in BP. Since part of the theoretical discussion between Modesto and the MTC proponents involved different grammaticality judgments of the relevant data, we have created an eye-tracking experiment that can indicate to us the degree of acceptability (or surprise) of inflected infinitives used in three distinct manners (i.e. having distinct antecedents) by BP speakers (college students in Rio de Janeiro). The experiment presented and discussed in section 3 below partially confirms Modesto's (2010, 2016a) claims that Brazilians are not surprised with inflected infinitives in complement clauses and have no problem assigning a control reading to a null category in subject position of an inflected nonfinite complement clause. The referential reading, in which a distant binder is taken as the antecedent of the null category, was thought to be ungrammatical; but the arbitrary interpretation was considered acceptable by Modesto. The examples below seem to indicate that Modesto was correct; however, when tested, referential readings were read faster than arbitrary readings, counter-intuitively:

- (1) a. Depois que as crianças foram encontradas,
 after that the children were found
 o casal percebeu terem feito besteira.
 the couple realized have.INF.PL done mistake
'After the children were found, the couple concluded that they made/had made a mistake.'

- b. Depois da prefeitura podar as árvores,

After the city.hall prune the trees

o casal percebeu terem feito besteira.

the couple realized have.INF.PL done mistake

'After the city hall pruned the trees, the couple concluded that they made/had made a mistake.'

Example (1a) has a preferable interpretation in which the couple realized that letting the children into the woods (for instance) was a mistake they made. That is the control interpretation. Would it be possible to have the interpretation in which the couple realized that the children made a mistake? That would be the referential interpretation. Finally, it would be possible to have an interpretation in which the couple realized that some arbitrarily defined group of people made a mistake. That reading is easier to get in (1b): after the city hall people pruned the trees we realized they had made a mistake; but it is also possible in (1a), even though it is not the preferred interpretation: after the children were found, the couple realized that someone made a mistake. We will come back to discussing these readings when analyzing the experimental data, in section 4.

The relevance of nonfinite inflection data was harshly criticized by Rodrigues and Hornstein (2013). For clarity, we provide here a concise recap of the dispute. To make his case against movement analyses of control, including the MTC, Modesto (2010) used examples of nonfinite inflection giving rise to partial control (PC) interpretations. Such constructions are assumedly rare in speech, but PC is not the only context of use of inflected infinitives in BP, and not the most common either. Other contexts of use are several kinds of adjunct clauses, subjective clauses and complement clauses of several classes of predicates and nominals (see MODESTO, 2016a for a fuller description). The number of contexts that either allow or require nonfinite inflection is so high as to make nonfinite inflection a salient characteristic of BP, as already noted by Lightfoot (1991, p. 99-102; a.o.). Rodrigues and Hornstein (2013, R&H from now on), on the other hand, consider the use of inflected infinitives to be “scarce” and not a feature of Brazilian’s internal grammar (or core

grammar), but something learnt in school.¹ However, concepts like core and periphrastic grammar are muddy, at best. One thing that could be considered periphrastic (if the concept could be made clear) in BP spoken grammar is the use of 3rd person accusative clitics, a kind of relic, not really used except when one is putting on airs or being very formal in their speech. Third person accusative clitics, which exist in modern EP, are normally not used in spoken BP, having been substituted by full (non-clitic) nominative pronouns or by null objects.² Although we Brazilians learn (in school or when reading literature) how to use those clitics, we almost never use them. When we do use them, they have a clear bookish ring to them. That is not the case with inflected infinitives. At all. Inflection in BP may not be a requirement for every speaker, but it definitely is a grammatical possibility for all speakers. Inflection (nonfinite and finite) is still an active ingredient of BP grammar, or in other words, still belongs to BP as a linguistic system (an I-language). In a sense, R&H equate nonfinite inflection to 3^p. accusative clitics, because both would be the product of schooling and reading. However, nonfinite inflection is nothing like accusative clitics in BP: whereas the use of nominative-like full pronouns in accusative position has become the social norm (see example (2a)), even among educated people, (2b) on the other hand is frowned upon, since it involves lack of subject-verb agreement, something that has been preserved in the speech of most BP speakers (especially in the big cities, by white people, etc.). It does not matter if the clause is finite

¹ R&H cite Pires and Rothman (2010) as saying that “the majority of Brazilian Portuguese speakers do not acquire inflected infinitives via an early native acquisition process but rather via late exposure to the standard dialect at school.” However, Rothman, Duarte, Pires and Santos (2013, p. 7) say textually that there are “standard BP speakers”: “inflected infinitives [...] are also found in adult BP, something expected if they correspond to a property of standard BP and if some families are speakers of the standard variety or switch between their colloquial dialect and standard BP (this might indeed explain the production of one inflected infinitive at 3;7 by A.C).” Although Rothman, Duarte, Pires and Santos do say that inflected infinitives are “limited” in the quote above, it is unclear in what sense. In the sense that only part of the population uses the paradigm of nonfinite inflection, then yes, the use of inflected infinitives is limited (to that population). However, the use of inflected infinitives (by that population that uses them) is not limited in any sense.

² See Cyrino and Lopes (2016) and Cyrino (2001) for description of usage and analyses of null objects in BP.

or nonfinite, BP seems to require subject-verb agreement. Therefore, Brazilians do hear infinitives inflecting all the time at work, in television, radio and the internet, whether they use the inflection themselves or not. Brazilian speakers normally accept as grammatical both versions (b-c) below, and react to them in different manners, given their sociolinguistic background (as discussed in section 2 below). More inflecting speakers will tend to accept (2c) better than (2b), less inflecting speakers will favor (2b) over (2c).

- (2) a. Eu vi ela ontem.
I saw her yesterday
'I saw her yesterday.'
- b. Eu fico contente de vocês ficar aqui.
I stay happy of you.PL stay.INF.Ø here.
'I am glad you guys stay here/are staying here.'
- c. Eu fico contente de vocês ficarem aqui.
I stay happy of you.PL stay.INF.PL here.
'I am glad you guys stay here/are staying here.'

R&H's argumentation departs from the fact that nonfinite inflection is not present in the speech of a number of Brazilian speakers, which would then make it a syntactic relic, similarly to third person accusative clitics. However, R&H depart from a wrong premise in supposing that BP, considered as an I-language, equals "popular BP", the norm spoken in rural areas of Brazil (see GUY, 1981).³ In reality, the fact that there are variants of BP in which nonfinite inflection has no phonological realization may be of no consequence for linguistic analysis. As seen above, (2b-c) have a complementary sociolinguistic behavior: there are speakers who tend to accept the variant with inflection (2c) better than the one without; and there are speakers, on the other hand, who tend to accept/produce the variant with no inflection (2b). Socio-political and historical facts in Brazil lead one to expect there to be a group of Brazilian speakers to which there would be no difference in grammaticality between

³ See also section 2, where the terms "popular" and "standard" are discussed.

(2b-c), just restrictions of sociolinguistic nature in their use (formality, etc.) related to the social norm. What matters is that, whether inflecting the verb or not, all BP speakers license overt subjects in the same nonfinite contexts (as shown in (2) and (3)), which suggests they have the same grammar. So, contrary to what R&H assumed, it is not the case that BP speakers have an internalized system with no nonfinite inflection; it is actually the case that all BP have nonfinite agreement, though some speakers have null or zero morphemes to mark such inflection. At least, that would be the simplest explanation for the licensing of overt subjects in contexts like (3b). The form in (3b) is socially stigmatized, which motivates the use of (3a) even in colloquial registers which, again, leads to the conclusion that sentences like (3a) are very common in BP. In fact, many sociolinguists are pointing out an increase in the use of (verbal and nominal) inflection in the last decades in Brazil (cf. LUCCHESI, 2012; OUSHIRO, 2015; SCHERRE; NARO, 2006).

- (3) a. Eu prefiro elas ficarem aqui comigo do que lá com o pai.
 I prefer they be.INF.PL here with.me of than there with the father
'I'd rather they stay here with me than with their father.'
- b. Eu prefiro elas ficar aqui comigo do que lá com o pai.
 I prefer they be.INF.Ø here with.me of than there with the father
'I'd rather they stay here with me than with their father.'

Overt inflection is not necessary in BP to license overt nominative nonfinite subjects, which means that all speakers (those who use overt inflection more than do not, and those who do not use overt inflection more frequently than they do) have the same internalized grammar, even though there is a lot of social variance. It is then paramount to investigate how inflection is actually used in BP. In what follows, we will be concerned with investigating which syntactic null category occupies the subject position of nonfinite inflected complement clauses in BP and EP. The possible analyses are described, in traditional generative-syntax terms, in (4) below. The *pro* analysis (4a) takes nonfinite inflection as able to license and identify a null pronominal with independent reference (as well as overt DPs). Therefore, the null subject position may refer to a distant (in terms of c-command) antecedent, or even to have no syntactic binder at all. This analysis seems to account for the behavior of complements of

propositional predicates such as *sugerir* (to suggest), *pensar* (to think), *acreditar* (to believe) in EP, (cf. RAPOSO, 1987; LANDAU, 2015; SHEEHAN, 2016) and the corresponding examples in (5).

The PRO analysis, in (4b), however PRO may be analyzed, takes subjects of inflected nonfinite verbs to be controlled (by normal rules of syntactic construal). Such analysis would be suitable to explain the interpretations of the null subject in the complement of desiderative predicates in EP (per SHEEHAN, 2016), and in the complement of any partial control predicate in BP (per MODESTO, 2010), see examples in (6). Regarding BP, the subject of an inflected nonfinite complement is either controlled (when null) or overt.⁴

Finally, the *pro*_{arb} analysis (4c) would explain the interpretation of null subjects with arbitrary interpretation, such as the ones in the examples in (7). According to R&H, the subject of nonfinite clauses in BP would be either like (4a) or (4c), whereas Modesto (2010, 2016a) argues that only (4b) and (4c) are possible in spoken BP. Our experiment shows that structure (4b) is definitely grammatical in BP. It also shows, surprisingly, that the referential reading is less problematic than the arbitrary reading. We will propose a syntactic explanation for that in section 4 below.

- (4) a. DP₁ control predicate [_{CP} *pro*₂ T-tense,+agr v]
 b. DP₁ control predicate [_{CP} PRO₁ T-tense,+agr v]
 c. DP₁ control predicate [_{CP} *pro*_{arb} T-tense,+agr v]
- (5) a. %A tua tia₁ sugeriu *pro*₂ encontrarmo-nos, mas sem ela₁ para atrapalhar. (EP)
 the your aunt suggested meet.INF.1PL REFL but without her to get.in.the.way
 ‘Your aunt suggested that we meet without her getting in the way.’
- b. Eu penso/lamento [terem (os deputados) trabalhado pouco]. (EP)
 I think/regret.1SG have.INF.3PL the deputies worked little
 ‘I believe/regret the deputies to have worked very little.’

⁴As pointed out by Sheehan 2016, there has been a growing understanding that, in many languages, overt subjects alternate with controlled null subjects in the same syntactic contexts (cf. LANDAU, 2000; SUNDARESAN; MCFADDEN, 2009; SUNDARESAN, 2014; MCFADDEN; SUNDARESAN, 2014). This is to be expected, since the controlled subject does receive Case (cf. LANDAU, 2006, 2008; SIGURÐSSON, 2008; BOBALJIK; LANDAU, 2009).

- (6) a. A tua tia₁ preferia PRO encontrarmo-nos₂ sozinhos. (EP)
 the your aunt preferred meet.INF.1PL REFL alone
 ‘Your aunt preferred that we met alone.’
- b. A presidente preferiu PRO nos encontrarmos outro dia. (BP)
 the president preferred REFL meet.INF.1PL another day
 ‘The president preferred for us to meet another day.’
- c. *A presidente preferiu PRO nos encontrarmos sem ela.⁵ (BP)
 the president preferred REFL meet.INF.1PL without her
 ‘The president preferred for us to meet without her.’
- (7) a. O Luiz₁ pensou de pro_{arb} terminarem o projeto sem ele₁. (BP)
 the Luiz thought about finish.INF.3PL the project without him
 ‘Luiz considered letting people finish the project without him.’
- b. O Luiz₁ pensou das meninas irem viajar (sem ele₁). (BP)
 the Luiz thought of.the girls go.INF.3PL travel.INF (without him)
 ‘Luiz thought about the girls travelling together (without him).’

BP cannot resort to structure (4a), that much is clear. That is logically deduced by the fact that null subjects may not have independent reference in BP even in finite clauses, so, obviously, nor in nonfinite clauses (cf. MOREIRA DA SILVA, 1984; FIGUEIREDO SILVA, 1994; DUARTE, 1995; NEGRÃO, 1999; GALVES, 1993, 2001; MODESTO, 2000a, 2000b; the collection of articles in ROBERTS; KATO, 1993; and KATO; NEGRÃO, 2000); so, the fact that R&H consider structure (2a) to be possible in BP is puzzling, since Rodrigues (2004) had also concluded that there is no referential *pro* in BP, in consonance with the studies listed above.

As for (4c), it is true that some inflected nonfinite clause’s subjects are interpreted arbitrarily, so Modesto (2010) treats (4c) as a possible structure. In the example (8a), below, for instance, the arbitrary reading is forced, since a control interpretation would cause a principle

⁵ What may occur is “A presidente preferiu nós nos encontrarmos sem ela”, with a full referential pronoun occupying the subject position of the nonfinite clause.

B violation (the pronoun in object position would be bound by the controlled subject). What Modesto (2010) claimed is that, in BP, in absence of any syntactic constrain, the first interpretation of (8b) is a control interpretation,⁶ not an arbitrary one (though that reading may be possible too, see CAVALCANTE, 2006; CAVALCANTE; DUARTE, 2009 on the use of arbitrary overt subjects in BP). To investigate this intuition, we set up an eye-tracking reading experiment in which BP speakers, after reading a clause like (8b) were asked to say who has proven the theorem. The results indicate that Modesto's intuitions were correct. Sentences like (8b) with a control interpretation are read faster and require less regressive eye movement than similar sentences with arbitrary interpretation.

- (8) a. O João acredita gostarem dele.
 the João believes like.INF.3PL of.him
 'João believes that people like him.'
- b. O João acredita terem provado o teorema.
 the João believes have.INF.3PL proved the theorem
 'João believes that they proved the theorem.'

In a recent addition to the debate, Sheehan (2016) shows that EP speakers may be divided between those who take the sentences in (6) to have the representation in (4a), showing obviation effects (cf. SITARIDOU, 2007), and those who use a control structure (4b) for the same sentences. The existence of such Brazilianized EP speakers could be the result of contact between the two varieties or it may indicate a general tendency of Portuguese in treating null subjects of inflected infinitives as controlled subjects. Another run of our eye-tracking experiment with EP speakers should shed some light on what syntactic category occupies the subject of nonfinite clauses for those speakers.

⁶ One of the anonymous reviewers, at this points, asks why there is no principle B violation here (in (8b)). The answer is contradictorily both simple and complicated. The answer is: because there is no pronoun there to violate principle B. What is there is a PRO (in that reading), not a pro. So now one would ask us what is PRO? That is the complicated part of the answer (see LANDAU, 2013, 2015).

Before turning to the experiment, the following section reviews some technical terms and concepts that are at the source of R&H misunderstanding, that BP is equal to Popular BP. Such a misconception leads them to three unwarranted claims: that inflected infinitives are scarce (in the speech of Brazilian speakers); that their use is regulated by normative grammar; and, that the paradigm of nonfinite inflection is not part of the internalized grammar of Brazilian speakers. The first claim is an empirical matter and interviews eliciting that kind of data are already underway at *Universidade Federal do Rio de Janeiro*, UFRJ (though a simple google search is already very telling). The second and third claims are obviously false (by speaker's intuition) and are tested (and disproven) in our experiment. Then, section 4 offers some theoretical discussion and interpretation of the data, as well as our concluding remarks.

2. What is BP after all?

What is any natural language? What kind of object is a language? The chomskyan tradition has answered those questions in these terms: a language L_1 is a mental state produced by the action of the social group (E-language) on the state L_0 of the mind, which then produces L_1 through a specialized *Language Acquisition Device* (CHOMSKY, 1965, p. 32). According to Chomsky in *Aspects*, in order to study a human language L , say BP, one should focus on “an ideal speaker-listener *in a completely homogeneous speech-community* who knows its language perfectly” (p. 3, our italics). The fact is that no such homogeneous community exists in Brazil, so the study of BP within generative syntax should be a careful one and unfounded assumptions like those in R&H should not be taken. As always, generative syntacticians should try to unveil the underlying competence of BP speakers that allows them to communicate in such seemingly chaotic linguistic context, as we try to do here.

It is not our intention to review the extensive sociolinguistic literature on (standard) BP and popular BP, but some inescapable concepts will have to be made clear, and some terminology has to be discussed, before any advance can be made in the generative-syntax description of BP (and, in that sense, we build upon GUY, 1981; TARALLO, 1988, 1993; LUCCHESI, 2001, 2004, 2012; LUCCHESI; BAXTER; RIBEIRO, 2009; NARO; SCHERRE, 2007; SCHERRE; NARO, 2006; MATTOS E SILVA, 2004; ILARI; BASSO, 2006; OUSHIRO, 2015;

FARACO, 2008, 2016). Most importantly, and almost consensually, the sociolinguistic situation in Brazil is described in most works just cited as being polarized between two norms: the norm of middle and upper class (urban white) Brazilians, usually called the “educated” norm (bad term, in our opinion), to which we have been referring as Standard BP; and the norm of working class and rural speakers, the “popular” norm. Such state of affairs, of course, has its roots in Brazilian history of slavery and colonialism. Brazil is believed to have imported more African slaves than any other country. An estimated 4.9 million slaves from Africa came to Brazil during the period from 1550 to 1875 (cf. <http://www.slavevoyages.org/>). According to Mattos e Silva, “starting from the 18th c. [...], especially in the urban areas that existed, there was a contrast between two possibilities: an Africanized Portuguese and a Europeanized one.” (MATTOS E SILVA, 2004, p. 21).⁷

From the 18th c. on, EP started to change (cf. GALVES; GALVES, 1995) and those changes solidified in 19th c. EP grammar, which was then adopted by early 20th c. Brazil as its “official language”, and, unfortunately, that grammar is still taught in schools in Brazil today. However, the BP spoken in the 18th c. did not change contiguously with EP. (White) BP kept being spoken in Brazil from the 18th c. on with its own characteristics, which culminated on a (white) Portuguese in Brazil today which is not identical to EP nor the Brazilian “official language”. Then, in the 20th c., the white-norm and the popular norm intermingled, education reached the lower classes and poor white immigrants (Europeans and Asians who came to Brazil fleeing poverty caused by World Wars I and II) ascended the social ladder “bringing to the heart of the cultured norm some of the structures of popular origin that they acquired in their initial contact with Portuguese.” (LUCCHESI; BAXTER; RIBEIRO, 2009, p. 53, our translation). All this has caused the 20th c. pickle Brazilians are in: we speak one language (with two norms) but we learn in our schools a different grammar, as if it were ours. The worst side of the problem is that the so called “educated” norm is more similar, in some respects, to EP, than the “popular” norm, which makes it harder for those speakers to learn the “official dialect” and grade well in university exams. One of such similarities between the “educated” norm and EP (normative grammar)

⁷ The original is: “a partir do século XVIII [...], sobretudo nas concentrações urbanas que já existiam, o embate se dava entre duas possibilidades: um português africanizado ou um português europeizado.” (MATTOS E SILVA, 2004, p. 21).

is the presence of more verbal and nominal inflection, although nonfinite inflection is used in radically different ways in EP and BP.

As sad as Brazilian social-linguistic situation is, it is inadmissible for us linguists to pretend that such differences do not exist and impose that BP I-grammar is the grammar of the popular speaker (in the sense of only considering data from popular speakers and dishing out all other data as being “peripheral”, “not core” features of BP grammar). It is also outrageous to imply that the I-grammar of “educated” speakers is regulated by “school-grammar”. As mentioned in the introduction, there are peripheral features in BP grammar, like 3p. accusative clitics, but inflected infinitives are completely different phenomena, with a large distribution of use, which is influenced by speakers’ intuitions (cf. MODESTO, 2016a).

As discussed by the authors cited above, there has been much influence in both directions, from “educated” to “popular” and from “popular” to “educated” norms, so we prefer to avoid those terms and urge others to avoid them. Standard and Popular BP are the most used terms in generative circles (again, to refer to the norm spoken by one or another BP speaker, so not to be confused with the “official” language we learn in school). Modesto (2016a) does not use the terms Standard and Popular to avoid misunderstandings and for seeing some residual racism on the term “popular”: he then refers to (more) inflecting speakers and less-inflecting speakers; terms that highlight the fact that Popular speakers do inflect sometimes (even infinitives) and that Standard BP speakers sometimes do not inflect (even in finite contexts). Whatever term one may choose, it must be clear that BP data may not be taken from one norm only, because the internalized grammar of Brazilian speakers is acquired within such a polarized sociolinguistic context. Assuming that the inflected infinitive has been “lost” in BP is just bad linguistics.

Given the history of Brazil, and according to the cited authors, actual BP is a continuation of Classical Portuguese, not a new (creole) language and, therefore, the null hypothesis is that nonfinite inflection is still part of BP (*qua* linguistic system). The study of nonfinite inflection in BP is interesting exactly because BP has lost the property usually referred to as “rich agreement”, the property of licensing a null pronoun by force of inflection alone (cf. MODESTO, 2000, among others). The logic behind Modesto’s (2010) analysis was to ask what happens in a situation where the colonizer speaks an Italian-type pro-drop language and the other half (at least) of the population speaks a probably creolized version of that language? Either the language becomes a non-pro-drop

one (in a way, like French did), or it becomes a discourse null-subject language, which is what has happened in Brazil (cf. PONTES, 1987; NEGRÃO; VIOTTI, 2000; MODESTO, 2008). What would happen in nonfinite contexts in the same social linguistic situation? It is highly plausible that null subjects in those contexts would become controlled, since null pronominal subjects are not licensed anymore in the language. Modesto's intuitions, plus collected data, indicated that this is what happened. And, as discussed in the following section, the passage from a pro-drop context to a control context when inflection becomes "weak" is also what is expected by the psycholinguistics literature.

To conclude this section, we believe that the (syntactic) contexts of use of nonfinite inflection in Brazil today and the interpretations of such structures are highly scientific interesting questions, considering Brazilian history and the synchronic social-political situation. We draw attention to the fact that the concepts of controller and antecedent and control and coreference are kept separate here. As discussed in the next section, when different antecedents are chosen as the reference of null categories, the control relation is a special one, for the null category can only have the closest argument as its antecedent in this relation. The notion of 'controller' (as closest antecedent considering c-command) is used in the psycholinguistic literature in exact the same sense it has in generative circles (cf. NICOL; SWINNEY, 1989; BETANCORT; MESEGUER; CARREIRAS, 2004; BETANCORT; CARREIRAS; ACUÑA-FARINA, 2006). The notion of coreference is different than the relation of control. If a null category can be coreferential with several different antecedents, it is most likely not a controlled position. What our experiment shows is that the subject position of nonfinite clauses in BP (inflected or noninflected) *is* a controlled position in BP when the subject is null. The phenomenon is not 'coreference' of the null category with the matrix subject, it is control, since the more distant antecedent is much harder to process. The experiment below is exactly tailored to show what kind of antecedents are in fact taken by null subjects of inflected infinitives. It shows that most speakers in most contexts tend to take the closest antecedent possible (other antecedents costing more, psychologically). It is therefore concluded that the category occupying the null subject position of an inflected infinitive in BP is PRO, not pro! The presence of PRO in inflected nonfinite clauses in BP clearly shows that the MTC is the wrong theory to use when dealing with Control (in any language). The corollary is that it cannot be used to explain *finite* null subjects in BP either.

3 The eye-tracking study

The processing of empty categories in sentence comprehension has been extensively investigated in Psycholinguistics. An important question in this literature has been the assessment of how and when coreferential relations are established between different types of empty categories and their antecedents. Nicol and Swinney (1989) is a seminal paper which examines the role of syntactic constraints on the reactivation and assignment of antecedents to explicit and implicit anaphoric elements during sentence comprehension. The authors review several studies which investigate coreferential assignment between different types of empty categories and their antecedents, concluding that “one consistent finding of recent studies is that reference-dependent sentential elements appear to cause reactivation of the antecedent noun phrases to which they refer” (cf. NICOL; SWINNEY, 1989, p. 6). McElree and Bever (1989) explore further the processing differences between different types of empty categories and conclude that “NP-movement gaps appear to activate their antecedent to a greater extent than PRO gaps” (cf. McELREE; BEVER 1989, p. 34). Fodor (1989) argues that the fact that empty categories of the PRO type are base-generated, that is, do not involve a moved element, makes them especially interesting, since there may be cases in which there are no antecedent warning the empty element before it is encountered.

Frazier, Clifton Jr. and Randall (1983) report a series of experiments testing the processing of empty categories of the PRO type in which they manipulate the main verb of the sentences in terms of their properties of requiring subject or object control.⁸ They measured the reading times for sentences such as (9) and (10) below:

- (9) Everyone liked the woman who the little child started PRO to sing those stupid French songs for __ last Christmas.
- (10) Everyone liked the woman who the little child forced PRO to sing those stupid French songs last Christmas.

They found that sentences with subject control verbs, such as (9) were read significantly faster than sentences with object control verbs

⁸ Frazier, Clifton and Randall (1983) also manipulated ambiguity as an additional factor in the study.

as (10). Based on this study the authors proposed the Most Recent Filler Strategy, which postulates that the human sentence processing mechanism prefers the most recent potential filler NP to fill the empty position of the infinitive verb, which is the processing version of syntactic principles like Rosenbaum's (1967) Minimal Distance Principle.

Osterhout and Nicol (1988), on the other hand, investigated the reactivation properties of PRO in a series of priming studies⁹ in sentences as exemplified below, that were composed of a matrix clause containing two potential antecedents and an infinitival clause:

- (11) The actress invited the dentist_i from the new medical center PRO_i to go to the party at the mayor's house.
- (12) The actress_i was invited by the dentist from the new medical center PRO_i to go to the party at the mayor's house.

In (11), the antecedent of PRO is 'dentist', a recent antecedent; in (12), it is 'actress', a distant antecedent. In their study, Osterhout and Nicol manipulated the position in the sentence in which a prime was launched and they found that when the prime was launched in the area of the last PPs in the sentences, only the actual antecedents were significantly activated: dentist in sentence (11), and actress in sentence (12). Based on these results they argue that the Most Recent Filler Strategy is not invoked in the processing of PRO and that "these results support the hypothesis that all structurally appropriate referents are reactivated" (cf. NICOL; SWINNEY, 1989, p. 17). Despite the controversy about the Most Recent Filler Strategy, both Frazier, Clifton and Randall (1983) and Osterhout and Nicol (1988), as well as McElree and Bever (1989) provide experimental evidence in favor of the so called "psychological reality" of PRO: when encountering PRO, a search for an antecedent in the sentence is activated. Using a priming technique, McElree and Bever (1989) show that the activation properties of PRO are observable, but they are not so strong as the activation properties of NP or Wh-traces. Frazier, Clifton and Randall (1983) measure reading times

⁹ These priming experiments consisted basically in the oral presentation of sentences, launching a written probe word or *prime* at specific points of the sentences. Subjects were asked to decide whether or not the probe word was in the sentence. Differences in accuracy rates and average reaction times in probe recognition in different areas of the sentence are then taken to indicate reactivation properties of referents in these areas.

during the reading of subject and object control sentences and propose that PRO prefers the most recent antecedent. Finally, Osterhout and Nicol (1988) argue that not necessarily the most recent, but the structurally appropriate antecedent of PRO is activated. In addition to the self-paced reading and priming techniques, Betancort, Meseguer and Carreiras (2004) study controlled PROs in Spanish, using eye-movement technology. The authors argue that the results of their two eye-tracking experiments show that readers do not seem to delay the selection of an antecedent of PRO. They conclude that “(...) when readers arrive at the empty category PRO, they begin a search for the antecedent of PRO. This search seems to be carried out quickly...” (cf. BETANCORT; MESEGUER; CARREIRAS, 2004, p.115). Betancort, Carreiras and Acuña-Farina (2006) further manipulate subject and object control verbs as well as adverbial clauses in Spanish in two eye-tracking experiments to show that “coreferential obligatory control is processed fast because it is launched from the same lexical platform that launches all fast-syntactic connections” (cf. BETANCORT; MESEGUER; CARREIRAS, 2004, p. 218).

Empty categories have been a productively investigated topic within the framework of two-stage models of processing, which discriminate between an earlier phase of syntactic parsing and a later phase, where nonstructural information is integrated. Two stage models as Garden Path Theory (GPT) (FRAZIER, 1979; FRAZIER; RAYNER, 1982), among several others) have proposed that the computation of gaps may be resolved as a first resort strategy (cf. FODOR, 1989), prioritizing the computation of grammatical structure over the integration of inferential and contextual information, which would only be accessed at a later integrative phase of comprehension. The present study presupposes the “syntax-first” framework of GPT and builds on the evidence that controlled PRO is psychologically real, triggering an antecedent search in the sentence as soon as it is encountered, by comparing the processing of controlled PRO with the processing of arbitrary PRO in inflected nonfinite clauses in BP. The main questions we wanted to address with our experiment were: (a) Is the search for an antecedent in the sentence the default process in comprehension when there is ambiguity between a controlled and an arbitrary interpretation for PRO? (b) To what extent is the arbitrary interpretation reading of PRO favored, when there is agreement mismatch between a local antecedent for PRO and the verb. (c) To what extent, if any, a referential interpretation of the null nonfinite subject is possible.

This study monitored the eye gaze of subjects (who were educated, but not necessarily inflecting speakers of BP) as they read sentences in BP made up by a contextualizing subordinate clause, a main clause and a final clause containing an inflected nonfinite clause, on a computer screen. The aim of the experiment was twofold. Firstly, we wanted to investigate the effect of obligatory control contexts on the processing of the sentences *vis à vis* contexts favoring indefinite, arbitrary and referential readings. Secondly, we manipulated number agreement between the subject and the inflected infinitive verb in order to assess whether the singular agreement would totally block a control reading of the sentence or whether a partial control interpretation would still be allowed.

We hypothesized that the latencies of first pass reading measures, the Total Fixation Duration (TFD) would be longer in the noncontrol conditions than in the control conditions, reflecting the higher processing cost of not establishing a local control syntactic relation which should be the default preference in the computation of the sentence. In line with the hypothesis that local control relation should be preferred as default in sentence processing, we also expected that latencies of regressive measures to the contextualizing subordinate sentences should be longer in the noncontrol interpretation conditions than in the control conditions, reflecting the higher cost of inferential processes *vis à vis* the computation of local grammatical relations. We also predicted that off-line interpretation questions, which were always probing for the referent of PRO in the inflected infinitive clause, should indicate preference for the control interpretation in control conditions in contrast to preference for the noncontrol inferential interpretation in noncontrol conditions, since the off-line interpretation measure should be able to capture post-syntactic integrative processes.¹⁰

Method

Participants

40 undergraduate students at the Federal University of Rio de Janeiro participated in the experiment voluntarily for course credit. They

¹⁰ In addition to TFD, the Fixation Count (FC) metric was also measured for each of the relevant areas across the conditions, but did not yield any robust effects and is not reported. Likewise, for the off-line dependent variable, significant effect were found only for the accuracy rates in the interpretation questions, decision times not being informative.

were all native speakers of BP with normal or corrected vision and without any history of reading problems.

Materials and Design

A total of 16 sets of experimental sentences like those displayed in Table 1 were created. Each sentence was immediately followed by a respective interpretation question probing for the referent of PRO/pro in the Inflected Infinitive clause. Examples of the corresponding questions for the sentences exemplified in Table 1 are displayed in Table 2. All experimental sentences were composed by an initial subordinate clause followed by a main clause and by a [PRO INFLECTED INFINITIVE CLAUSE]. There were two independent variables crossing in a 2x2 design, namely, Control context, which had two levels, obligatory control (O) or no control (N), and Number, which could be singular (S) or plural (P). The crossing of these two independent variables, each with two levels, generated four conditions: No-control plural (NP), Obligatory control plural (OP), No control singular (NS), Obligatory control singular (OS). The no control conditions were additionally manipulated in terms of two sub-conditions: 8 sentences had a referential DP in the subordinate clause which could potentially be the controller of the PRO in the Inflected infinitive clause and other 8 sentences did not include such a referential DP in the subordinate clause, forcing an exclusively arbitrary interpretation for the PRO.

Obligatory-control sentences always contained, in the initial subordinate clause, a verb whose cataphoric subject was the subject DP in the main clause. This DP was a semantically plausible subject for all the verbs in the sentence: the subordinate clause verb, the main clause verb and the inflected nonfinite clause verb. The non-control sentences, on the other hand, never allowed the possibility that the main clause subject DP could also be the cataphoric subject of the preceding subordinate clause verb. Non-control sentences always had a DP subject in the main clause which was not likely to be the antecedent for the PRO in the nonfinite clause. In the adjunct clause, non-control sentences either had no semantically plausible controller DP or there was a DP which could plausibly be a distant antecedent for the PRO/pro in the nonfinite clause. The independent variable 'number' included a level in which the DP subject and the main clause verb agreed in the plural (P) and a level in which the DP subject and the main clause verb agreed in the singular (S).

Adjunct clause verbs as well as inflected nonfinite verbs were always in the plural. Experimental sentences were distributed in a Latin square design, generating four versions of the experiment. In addition, another 32 sentences with different types of structures were created to serve as fillers. Both for experimental and distractor items, sentence length was controlled to be within 25 to 30 metric syllables range. The 48 sentences (16 experimental and 32 fillers) in each version were pseudo-randomized to be presented in a fixed random order interspersing distracting sentences among the experimental sentences, and guaranteeing that the first and the last two sentences were never experimental sentences.

TABLE 1 – Examples of sentences used in the eye-tracking experiment

Cond		Adjunct clause	Main clause	Inflected nonfinite clause
NP	a r b	Como os feridos foram achados logo <i>As the wounded were found soon</i>	os repórteres julgaram <i>the reporters judged</i>	terem salvo muitas vidas <i>to have saved many lives</i>
	r e f	Só quando os bebês foram examinados <i>Only when the babies were examined</i>	os cuidadores perceberam <i>the caretakers realized</i>	terem sujado as fraldas <i>to have dirtied the diapers</i>
OP		Como chegaram logo ao local da queda <i>As (they) arrived at once in the crash area</i>	os bombeiros julgaram <i>the firefighters judged</i>	terem salvo muitas vidas <i>to have saved many lives</i>
NS	a r b	Como os feridos foram achados logo <i>As the wounded were found soon</i>	o repórter julgou <i>the reporter judged</i>	terem salvo muitas vidas <i>to have saved many lives</i>
	r e f	Só quando os bebês foram examinados <i>Only when the babies were examined</i>	o cuidador percebeu <i>the caretaker realized</i>	terem sujado as fraldas <i>to have dirtied the diapers</i>
OS		Como chegaram logo ao local da queda <i>As(they) arrived soon in the crash area</i>	o bombeiro julgou <i>the firefighter judged</i>	terem salvo muitas vidas <i>to have saved many lives</i>

TABLE 2 – Examples of interpretation questions after presentation of each sentence

Cond		Interpretation questions	
NP	arb	Quem salvou muitas vidas? <i>Who saved many lives?</i>	(A) os repórteres (A) <i>the reporters</i> (B) outras pessoas (B) <i>other people</i>
	ref	Quem sujou as fraldas? <i>Who dirtied the diapers?</i>	(A) os cuidadores (A) <i>the caretakers</i> (B) outras pessoas (B) <i>other people</i>
OP		Quem salvou muitas vidas? <i>Who saved many lives?</i>	(A) os bombeiros (A) <i>the firefighters</i> (B) outras pessoas (B) <i>other people</i>
NS	arb	Quem salvou muitas vidas? <i>Who saved many lives?</i>	(A) o repórter (A) <i>the reporter</i> (B) outras pessoas (B) <i>other people</i>
	ref	Quem sujou as fraldas? <i>Who dirtied the diapers?</i>	(A) o cuidador (A) <i>the caretaker</i> (B) outras pessoas (B) <i>other people</i>
OS		Quem salvou muitas vidas? <i>Who saved many lives?</i>	(A) o bombeiro (A) <i>the firefighter</i> (B) outras pessoas (B) <i>other people</i>

Procedures

A TOBII TX300 eye-tracker monitored subject's eye-movements. Participants viewed the stimuli binocularly on a monitor 65 cm from their eyes. Sentences were written in Courier New 28pt Font (True Type) and were displayed in a single line. Before the experiment started, subjects read a set of instructions on how to perform the experiment. The instructions told them to read the sentences at a very fast rate, but trying to comprehend the text as well as they could, as they would be asked an interpretation question at the end of each sentence. By pressing the space bar on the keyboard, subjects self-monitored the duration of each sentence on the screen. In order to guarantee a fast reading of the sentences, there was a timeout of 2 seconds, after which the sentence was replaced on the screen by the interpretation question. Participants were asked to try to read the sentence as fast as they could and press the keyboard bar before the timeout. Therefore, after reading the sentence, subjects typically would press the bar before the timeout, in order to call another screen which contained a question with two possible answers, each preceded by (A) and (B) and subjects were instructed to press an (A) or a (B) button in the keyboard. If subjects delayed pressing the bar,

the screen with the interpretation question would appear automatically after two seconds. There was a 5 second timeout for the interpretation question screen. The experimenter then started a calibration session which consisted in having the subject follow a red ball moving around the screen. If the calibration was successful as indicated by a screen after the calibration session, the experimenter would start a practice session, otherwise a recalibration session would take place. The practice session consisted of three sentences and was observed by the experimenter. If the subject confirmed that he had understood the instructions well and could do the experiment, the instructor would start the experiment and leave the room. A full session would generally last for 25 to 30 minutes.

Results

For the purpose of the analysis, the texts were segmented in three areas as shown in Table 1: adjunct clause, main clause and inflected nonfinite clause. Results for two different eye-movement measures which displayed robust effects will be reported: (1) Total Fixation Duration (TFD), which refers to the summed durations of all fixations made on each of the three areas and on the whole sentence, including later fixations resulting from regressive movements from subsequent words in the sentence or re-reading the sentence starting at words prior to the target; (2) Second pass duration on the subordinate clause area, which refer to the sum of all refixations directed to that area, either from the left or from the right. The observed differences were statistically evaluated by an ANOVA for subjects (F1). TFDs are indicated in Table 3:

TABLE 3 – TFDs in milliseconds

Conditions	Adjunct clause	Main clause	Inflected nonfinite clause	TOTAL
NP	860	418	487	1765
OP	692	332	435	1459
NS	754	398	474	1626
OS	722	376	452	1550

The 2x2 ANOVA for subjects crossing the two within subjects factors control and number carried out for the sum of TFDs of each condition shows a highly significant main effect of control ($F(1, 159) = 213, p < 0,000001$), a marginally significant main effect of number ($F(1,159) = 3,73 p < 0,055374$) and a highly significant interaction of control*number ($F(1,159) = 98,2 p < 0,000001$). Pairwise comparisons between the conditions were also performed yielding highly significant results for all relevant crossings: NP x OP ($t(159)=13,13 p < 0,0001$), NS x OS ($t(159)=9,13 p < 0,0001$), NP x NS ($t(159)=9,14 p < 0,0001$) and OP x OS ($t(159)=4,79 p < 0,0001$).

The area of the inflected nonfinite clause was also submitted to a 2x2 ANOVA by subjects, showing a highly significant main effect of control ($F(1,159) = 83,0 p < 0,000001$), a nonsignificant main effect of number ($F(1,159) = 0,218 p < 0,641306$), and a significant interaction of control*number ($F(1,159) = 11,1 p < 0,001081$). Pairwise comparisons yielded significant results for NP x OP ($t(159)=7,74 p < 0,0001$), NS x OS ($t(159)=4,30 p < 0,0001$), OP x OS ($t(159)=3,30 p < 0,0012$), but there was no significance in the comparison between NP x NS ($t(159)=1,74 p < 0,0832$).

An ANOVA by subjects was also carried out in the region of the main clause and showed a highly significant main effect of control ($F(1,159) = 190 p < 0,000001$), a significant main effect of number ($F(1,159) = 9,53 p < 0,002390$) and a highly significant interaction of control*number ($F(1,159) = 74,4 p < 0,000001$). Pairwise comparisons yielded significant results for NP x OP ($t(159)=13,89 p < 0,0001$), NS x OS ($t(159)=4,93 p < 0,0001$), NP x NS ($t(159)=3,24 p < 0,0014$) and OP x OS ($t(159)=11,04 p < 0,0001$).

First and Second pass fixation durations on the subordinate clause area were also obtained for all conditions and are shown in Table 4.

TABLE 4 – Second pass fixation durations on subordinate clause

Condition	1 st pass	2 nd pass	Total
NP	560	300	860
OP	558	134	692
NS	566	188	754
OS	556	166	722

First pass fixation durations did not exhibit any significant differences in 2x2 ANOVA by subjects neither for main effect of control ($F(1,159) = 1,27$ $p < 0,261362$), nor for main effect of number ($F(1,159) = 0,034$ $p < 0,853785$). No significant interaction was observed between the factors ($F(1,159) = 0,641$ $p < 0,424381$). As expected, there was no difference in pairwise t-tests between the conditions.

Second pass fixation durations on the subordinate clause, on the other hand, displayed significant differences in the 2x2 ANOVA by subjects. There was a highly significant main effect of control ($F(1,159) = 2570$ $p < 0,000001$) and a highly significant main effect of number ($F(1,159) = 539$ $p < 0,000001$). Pairwise comparisons were also significant for all relevant crossings, in the expected directions. NP x OP ($t(159) = 44,54$ $p < 0,0001$), NS x OS ($t(159) = 10,61$ $p < 0,0001$), NP x NS ($t(159) = 32,40$ $p < 0,0001$) and OP x OS ($t(159) = 13,45$ $p < 0,0001$)

A subanalysis comparing the two sub-conditions of the no-control condition was also performed to establish any possible effects of referentiality. As explained in the Materials & Design section above, 8 sentences had a referential DP in the subordinate clause which could potentially be the controller of the PRO in the final clause and other 8 sentences did not include such a referential DP in the subordinate clause, forcing an arbitrary interpretation for the PRO. TFDs for the whole sentences, considering these sub conditions of the no control conditions are indicated in Table 5.

TABLE 5 – TFDs for sub conditions of N condition

NP REF	NP ARB	NS REF	NS ARB
1724	1806	1584	1667

The subject 2x2 ANOVA showed a highly significant referentiality effect ($F(1,79) = 29,3$ $p < 0,000001$) and a highly significant main effect of number ($F(1,79) = 62,4$ $p < 0,000001$), but no interaction between the two factors ($F(1,79) = 0,001$ $p < 0,996830$). The relevant pairwise t-tests also yielded significant results in the expected directions. NS REF x NS ARB indicated that singular sentences with a potential controller referential DP in the adjunct clause were read significantly faster than singular sentences which did not include such a potential controller referential DP ($t(79) = 5,63$ $p < 0,0001$). Likewise, NP REF x NP ARB

indicated that REF plural sentences were significantly faster than ARB plural sentences ($t(79)=3,48$ $p<0,0008$). NP REF x NS REF also showed a significant difference ($t(79)=6,64$ $p<0,0001$), as well as NP ARB x NS ARB ($t(79)=6,26$ $p<0,0001$).

Off-line interpretation questions were computed and expressed in percentages of CONTROL choices¹¹. Table 6 indicates the results obtained for all conditions, including the N conditions referentiality manipulations in arbitrary (arb) or referential (ref) interpretations:

TABLE 6 – Off-line choices for control interpretation (%)

Cond	NP	NP	NS	NS	OP	OS
REF	arb	ref	arb	ref	-----	-----
Cont. INT	36%	28%	29%	24%	83%	62%

Chi-square analyses were performed, yielding the following results: OP x OS ($X^2= 6,08$, $p= 0,01$); OP x NP arb ($X^2=37,1$, $p= 0,0001^{***}$); OP x NP ref ($X^2= 54,5$, $p= 0,0001^{***}$); NP arb x NP ref ($X^2= 2,0$, $p= 0,15^{ns}$); OS x NS arb ($X^2= 23,9$, $p= 0,0001^{***}$); OS x NS ref ($X^2= 33,5$, $p= 0,0001^{***}$); NS arb x NS ref ($X^2=0,94$, $p=0,33^{ns}$); NP arb x NS arb ($X^2= 1,5$, $p= 0,21^{ns}$); NP ref x NS ref ($X^2= 0,61$, $p=0,43^{ns}$).

Discussion

The effects obtained for the sum of TFDs of each condition showed that readers took longer to read the noncontrol conditions than to read the control conditions. This overall result already gives us a first answer to our question about the default preference when comparing the processing of controlled PRO with the processing of arbitrary PRO, in inflected nonfinite clauses in BP. When it is possible to establish a local control syntactic relation, the sentence is processed faster than in the indefinite or arbitrary reading cases. In line with Betancort, Meseguer and Carreiras (2004), when readers arrive at the empty category PRO, they begin a search for the antecedent of PRO. The simplest search of

¹¹ Average response times were not significantly different in any of the comparisons between conditions and are not reported.

all takes place in the OP condition in which the subject DP in the main clause is both the cataphoric subject of the subordinate clause verb and of the inflected infinitive verb. Accordingly, the OP condition receives the highest rate of control interpretation when compared to all other conditions. The second fastest reading is OS, the PC condition, in which there is a number mismatch between the main clause subject and both the verbs of the subordinate clause and of the inflected infinitive clause. In this condition, the semantic/pragmatic fit between the main clause subject DP and the inflected infinitive verb allowed for the partial control of the PRO, as established in the off-line measure, which indicated the second highest rate of control interpretation. Again, having a local controller, even if partial, facilitated the PRO coreference search, confirming the preference for a local controller of the PRO as the default process in the resolution of nonfinite clauses. Noncontrol sentences took longer to read than Control or PC sentences, receiving lower rates of control interpretation than control or partial control sentences, as expected.¹²

Total fixation durations of NS sentences are longer than OS sentences and OP sentences, but shorter than NP sentences. Both NP and NS sentences presented a subject DP in the main clause which was not a plausible controller for PRO in the inflected nonfinite clause, opening the way for an arbitrary interpretation. The number mismatch between the subject DP in the main clause and the inflected nonfinite verb should contribute to additionally rule out the control analysis, pushing the reader even more towards the arbitrary analysis. NP sentences, in contrast, establishing plural agreement matching between the subject DP of the main clause and the plural agreement feature of the inflected infinitive verb would create a potential ambiguity between an arbitrary and a control reading for the PRO, leading the reader to fixate longer the sentences in this condition.

As indicated in the last column of Table 3, the total reading durations per condition were distributed as follows: OP < OS < NS <

¹² Note that in the N conditions there is still a 24% to 36% choice for the control interpretation. As indicated in the statistical analyses, these differences are not significant across the N conditions, but the very fact that even in the N conditions a control interpretation is residually entertained can be taken as a further evidence of the default control analysis proposed in this paper.

NP. Having a local controller for the PRO, OP sentences are the easiest to read, as illustrated by the heatmap¹³ of a typical reading, in Figure 1.

FIGURE 1 – OP sentence:

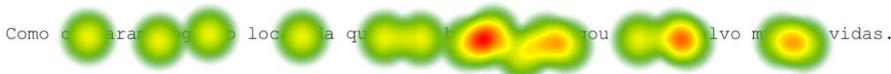
*Como chegaram logo ao local da queda, os bombeiros julgaram
terem salvo muitas vidas*
As they arrived at once in the crash area, the firefighters judged
to have saved many lives.



OS sentences take longer to read than OP sentences, as there is a number mismatch between the felicitous subject DP of the main clause and the inflected infinitive verb. However, partial control is still possible and they are resolved in shorter reading times than the N sentences, which were intended to block the local control interpretation and require more exhausting visual search for a controller for the PRO. Not finding a possible local controller in an N sentence, readers should invoke the arbitrary reading as a last resort. Figure 2 provides the heatmap obtained during the reading of an OS sentence and Figure 3 exemplifies the reading of an NS sentence.

FIGURE 2 – OS sentence:

Como chegaram logo ao local da queda, o bombeiro julgou terem salvo muitas vidas.
As they arrived at once in the crash area, the firefighter judged to have saved
many lives.



¹³ A heatmap is a graphical representation of data where the individual values contained in a matrix are represented as colors. *The areas where readers looked the most are colored red*; the yellow areas indicate fewer views, followed by the least-viewed green areas.

FIGURE 3 – NS ref sentence:

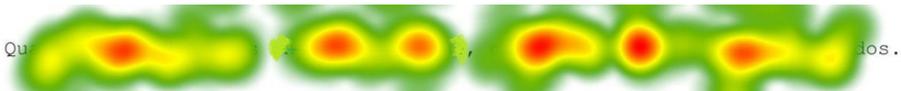
Só quando os bebês foram examinados, o cuidador percebeu terem sujado as fraldas.
 Only when the babies were examined, the caretaker realized (them) to have dirtied the diapers.



Notice that in Figure 3, the correct interpretation of the null category requires an inferential chain, since in addition to number mismatch, it would not be plausible to analyze the DP *o cuidador* “the caretaker” as the agent of having dirtied the diapers. In the NSref condition there is a possible controller in the adjunct clause, but not in the NS arb condition exemplified in Figure 4, in which, in the adjunct clause, there is no possible antecedent. Notice that, in Figure 4, the correct interpretation of PRO requires an inferential chain, since in addition to number mismatch, it would not be plausible to analyze the DP *the candidate* as the agent of being ill intentioned. Additionally, in the subordinate clause there is no possible controller for PRO either. The arbitrary interpretation is invoked only after the first resort search for a local controller is unsuccessful, resulting in even longer fixations during reading.

FIGURE 4 – NS arb sentence:

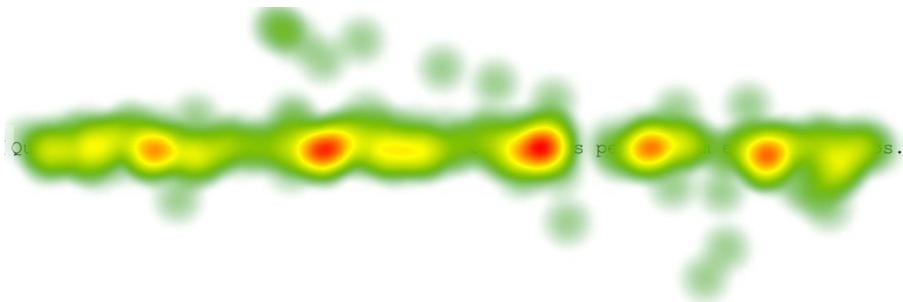
Quando o concurso foi cancelado, o candidato julgou estarem mal-intencionados.
 When the selection was cancelled, the candidate judged (them) to be ill intentioned.



In the NP condition, agreement between the main clause subject DP and the inflected infinitive verb seems to make the search for a local controller even longer, possibly misleadingly allowing for the initial search for a local controller which is then discarded because it is not a felicitous agent, even though matching in formal features. The search proceeds then to find a possible controller in other areas of the sentence. Figure 5 illustrates the heatmap of an NP ref sentence.

FIGURE 5 – NP ref sentence:

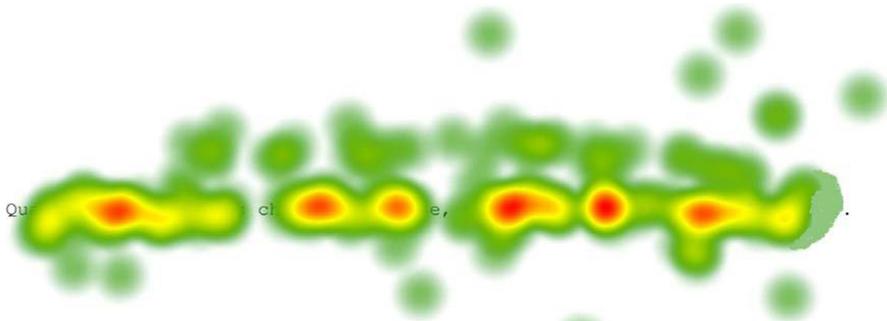
Quando os terroristas chegaram na base, os pilotos perceberam estarem armados.
When the terrorists arrived at the base, the pilots realized (them) to be armed.



Finally, Figure 6 illustrates the heatmap of an NP arb sentence, which requires the longest average fixations of all conditions, since there is no felicitous DP in the sentence which could be the proper controller for the PRO. The arbitrary interpretation of the-PRO is the costliest.

FIGURE 6 – NP arb sentence:

Quando os dados chegaram na base, os cientistas perceberam estarem em órbita.
When the data arrived at the base, the scientists realized to be in orbit.



As shown in the Results section, comparative analyses of other areas of the sentences between the four conditions also display main effects of control. Pairwise comparisons between the conditions also indicated interesting significant differences in the areas of the main clause and of the inflected nonfinite clause in the expected direction, attesting that the noncontrol conditions are harder to process than the control conditions.

Table 4 shows comparative analyses between the total fixation durations during the first pass and the second pass in the area of the subordinate clause. As expected the first pass measures do not exhibit any significant differences between the conditions, since at this point readers had not been exposed to the independent variables. Notice, however, that second pass measures differ in the expected direction: readers exhibited significantly lesser regressions to the adjunct clause in the OP condition, which resolves the PRO coreference in a very straightforward manner than in any other condition. Actually, second pass measures in the subordinate clause confirm the hierarchy established for the TFD of the sentences: $OP < OS < NS < NP$, indicating that regressions are progressively needed to compute the sentences as they move away from local syntactic control to more arbitrary or ambiguous interpretations, again, in line with the hypothesis that local control relations should be preferred as default in sentence processing.

To conclude, the subanalyses entertained here between the arbitrary x referential manipulations of the N sentences also confirm the hypothesis. As shown in Table 5, referential noncontrol sentences are read faster than arbitrary noncontrol sentences, again indicating the preference for local syntactic relations which can resolve coreference more readily and are therefore preferred to inferential processes, as predicted by two-stage models of sentence processing, as Garden Path Theory.

4 Theoretical discussion and conclusions

As detailed in section 3, in line with several previous psycholinguistic studies, especially work on structural models of sentence processing, which have very consistently established the priority of structural computations over inferential interpretations, the present paper has demonstrated, among other things, the priority of structural processes over inferential interpretations in the processing of PRO in BP: when readers arrive at the empty category PRO, they begin a search for the antecedent of PRO as a default. The most accessible antecedent for the null category in our test-sentences was the subject of the matrix clause and when context and lexical choices allowed that control interpretation, reading of the sentence was easiest. This means that controlled PRO in inflected nonfinite clauses is not only psychologically real in BP, it is actually the preferred option in a task in which a strict comparison with arbitrary PRO is entertained.

This means that none of the problems raised by R&H are relevant and the data in Modesto (2010) does refute movement analyses of Control. All these assertions should be unpacked here, if we had the space. For now, let us just take notice of how the MTC makes the wrong predictions about BP, when actual data is taken into consideration. Besides the psycholinguistic discussion, we have seen, in section 2, that nonfinite inflection has been present in the speech of BP speakers (actively, for some speakers; passively for others) for at least three centuries

Given BP's sociolinguistic history and context, it is easy to see how Portuguese is spoken scarcely in the Brazil of the 16th c. (i.e. exclusively at or around the "big" cities that were formed then (Salvador, BA; Pernambuco, PE; Rio de Janeiro, RJ e São Paulo, SP). However, Portuguese becomes the only language acquired by most people born in Brazil from the 18th c. on. Such nativization process lasted only two centuries, speakers learnt their Portuguese from very different sources (white BP vs. black BP). Then in the 20th c., these two variants of BP were mixed and mingled. But there should be no problem since we have been using the same language, after all. It is unimaginable that through the 19th c. white speakers had a much different grammar than black speakers. It would have been impossible for BP to keep the romance-type pro-drop grammar brought here by the 17th and 18th c.-Portuguese people who immigrated, because some (roughly half) of the speakers (slaves and Indians) had no inflection in their system during that time. So already in the 19th c., at least, BP probably was already a non-pro-drop language (to every speaker; no matter how much inflection they used). It is possible that very few conserved the older system, or that some people were bilingual speakers of both varieties of BP at that time (see KROCH, 2001 on grammars in competition). But that is unlikely for nowadays BP, since there is no evidence for two different linguistic systems in Brazil, just a lot of morphological variation. If the conclusions expressed here for the discussion in section 2 are correct, i.e. if all Brazilians have nonfinite inflection whether using it overtly or not, this is very much in accordance with a third factor explanation (cf. CHOMSKY, 2005, as pointed out by one of the anonymous reviewers of this paper) for syntactic data of natural languages. Interestingly, the present work has brought about relevant data for a contemporary challenge in the relatively new domain of Experimental Syntax – the interplay between processing algorithms, Universal Grammar (UG) and the grammar of particular

languages. Principles of computational efficiency have been analyzed as third factor effects (cf. CHOMSKY, 2005). If the PRO control processing default shown in the present paper can be claimed to be an efficiency processing algorithm it raises interesting architecture issues concerning the interaction of processing and grammar. In this respect, we follow the proposal in Trotkze, Bader and Frazier (2013) that “properties of performance systems can play an important role within the biolinguistic perspective on language by providing third-factor explanations for crucial design features of human language” (TROTKZE; BADER; FRAZIER, 2013, p. 28). Of course, third factor effects must interact with UG principles and with grammatical properties of specific languages, sometimes bringing about tensions when they apply in the derivations, but they probably do not override grammar internal principles, as argued for in Di Sciullo and Aguero (2008) on independent grounds. We intend to keep on this line of research, incorporating EP experimental data and comparing the results with BP.

The problem with that explanation is that BP continues to be a language with null subjects, even if not a pro-drop language, using a different strategy that does not rely on inflection (like Chinese). Subjects of inflected clauses can still be null in BP, but they now have a controlled interpretation (in most contexts), as the experiment demonstrates. The odd result, however, is that our informants seem to have less trouble getting a non-controlled, referential reading of the null subject than getting an arbitrary reading. Though unproblematic for the psycholinguistics part, this result is tricky to explain syntactically. If BP inflection does not license null pronouns anymore, how are referential readings achieved by BP speakers? A mandatory quasi-topic position in BP, as proposed in Modesto (2008), nicely explains those readings, given a suitable context, as our test-sentences provided. Take (13a) as an example. Its representation would be (13b), with an elided topic:

- (13) a. Só quando os bebês foram examinados, o cuidador percebeu terem
 only when the babies were examined, the caretaker noticed have.INF.3PL
 sujado as fraldas
 soiled the diapers
‘When the babies were examined, the caretaker noticed that they have soiled the diapers.’

- b. [[Só quando os bebês foram examinados] [os-bebês,
[o cuidador percebeu [PRO_i [PRΘ_i terem sujado as fraldas]]]

The example in (14a), on the other hand, has the structure in (14b).

- (14) a. Como chegaram logo ao local da queda,
as arrive.PST.3PL soon to.the place of fall
o bombeiro julgou terem salvo muitas vidas.
the firefighter judged have.INF.3PL saved many lives
*'As(they) arrived soon in the crash area, the firefighter judged
to have saved many lives.'*
- b. [[Como chegaram logo ao local da queda] [o bombeiro,
[ec_i julgou [PRO_i [PRΘ_i terem salvo muitas vidas]]]]]

We believe that the empty category in (14c) marked as *ec* is just a minimal pronoun (as in SZABOLCSI, 2009), the same category that produces control (as in LANDAU, 2015), which gets the values for its features by being bound by “o bombeiro”, the firefighter, the closest possible antecedent. PRO itself is a minimal pronoun that climbs to the quasi-topic position, still finds no value for its person feature, which then causes the control phenomenon (all as described in LANDAU, 2015). It can be seen, then, that BP data involving inflected infinitives are unproblematic when the most plausible analysis of control is assumed.

Acknowledgments

This work is part of the State University of Campinas project “Portuguese in space and time: linguistic contact, grammars in competition and parametric change” (FAPESP 12/06078-9), per the first author and of the CNPq research project “Experimental Syntax and Semantics: timecourse and underspecification in the on-line implementation of grammar” (CNPq 302989/2015-3), per the second author. We both wish to thank our informants and the students involved in the experiment described here, as well as two anonymous reviewers who made valuable comments on our first draft. All remaining errors are ours, of course.

References

AMBAR, M. “Aux-to-COMP” and lexical restrictions on verb movement. In: CINQUE, G.; KOSTER, J.; POLLOCK, J. Y.; RIZZI, L.; ZANUTTINI, R. (Org.). *Paths towards universal grammar*. Amsterdam: John Benjamins, 1994. p. 1-24.

AMBAR, M. Inflected infinitives revisited: Genericity and single event. *Canadian Journal of Linguistics*, Cambridge University Press, v. 43, p. 5-36, 1998.

BETANCORT, M.; CARREIRAS, M.; ACUÑA-FARINA, C. Processing controlled PROs in Spanish. *Cognition*, Elsevier, v. 100, p. 217-282, 2006. <https://doi.org/10.1016/j.cognition.2005.04.001>

BETANCORT, M.; MESEGUER, E.; CARREIRAS, M. The empty category PRO: Processing what cannot be seen”. In: CARREIRAS, M.; CLIFTON, C. (Org.). *The online study of sentence comprehension: Eyetracking, ERP, and beyond*. C. Brighton: Psychology Press, 2004.

BOBALJIK, J.; LANDAU, I. Icelandic Control is not A-movement: The Case from Case. *Linguistic Inquiry*, MIT Press, v. 40, p. 113-132, 2009. <https://doi.org/10.1162/ling.2009.40.1.113>

BOECKX, C.; HORNSTEIN, N. The Virtues of Control as Movement. *Syntax*, Wiley Online Library, v. 9, p. 118-130, 2006.

BOECKX, C.; HORNSTEIN, N.; NUNES, J. Control as movement. Cambridge: Cambridge University Press, 2010. <https://doi.org/10.1017/CBO9780511761997>

CAVALCANTE, S. R. O. *O uso de SE com infinitivo na história do português: do Português Clássico ao Português Europeu e Brasileiro modernos*. 2006. Tese (Doutorado) – Unicamp, Campinas, 2006.

CAVALCANTE, S. R. O.; DUARTE, M. E. L. Sujeitos de referência arbitrária em sentenças infinitivas do português e o parâmetro do sujeito nulo. In: TSIPLAKOU S.; KARYOLEMOU, M.; PAVLOU, P. (Org.). *Language Variation – European perspectives II: selected papers from the 4th international conference on language variation in Europe*. Amsterdam: John Benjamins, 2009. p. 47-58.

CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge: The MIT Press, 1965.

CHOMSKY, N. Three factors in Language Design. *Linguistic Inquiry*, v. 36: n. 1, p. 1-22, Winter 2005. doi: 10.1162/0024389052993655

CYRINO, S. M. L. O objeto nulo no português do Brasil e no português de Portugal. *Revista da ABRALIN*, Abralín, v. 25, p. 173-181, 2001.

CYRINO, S. M. L.; LOPES, R. Null objects are ellipsis in Brazilian Portuguese. *The Linguistic Review*, De Gruyter, v. 33, p. 1-19, 2016.

DI SCIULLO, A.M.; AGUERO, C. The delay of Condition B Effect and its Absence in Certain Languages. *Language and Speech*, Sage Publications, v. 51, p. 77-100, 2008. <https://doi.org/10.1177/00238309080510010601>

DUARTE, M. E. *A perda do princípio “evite pronomes” no português brasileiro*. 1995. Tese (Doutorado) - Unicamp, Campinas, 1995.

FARACO, C. A. *Norma culta brasileira: desatando alguns nós*. São Paulo: Parábola, 2008.

FARACO, C. A. *História sociopolítica da língua portuguesa*. São Paulo: Parábola, 2016.

FIGUEIREDO SILVA, M. C. *La position sujet en Portugais Brésilien*. 1994. Dissertation (Doctoral) - Université de Genève, Genève, 1994.

FODOR, J. D. Empty categories in sentence processing. *Language and Cognitive Processes*, Taylor & Francis Online, v. 4, p. 155-209, 1989.

FRAZIER, L. *On comprehending sentences: syntactic parsing strategies*. 1979. Dissertation (Doctoral) - University of Massachusetts, Connecticut, 1979. [Reproduced by Indiana University Linguistics Club.]

FRAZIER, L.; CLIFTON JR., C.; RANDALL, J. Filling gaps: decision principles and structure in sentence comprehension. *Cognition*, Elsevier, v. 13, p. 187-222, 1983. [https://doi.org/10.1016/0010-0277\(83\)90022-7](https://doi.org/10.1016/0010-0277(83)90022-7)

FRAZIER, L.; RAYNER, K. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, Elsevier, v. 14, p. 178-210, 1982. [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1)

GALVES, A.; GALVES, C. *A case study of prosody driven language change: from classical to modern EP*. 1995. Ms. - UNICAMP-USP, Campinas, 1995. Disponível em: <http://www.lilf.uam.es/~clase/acceso_local/prosodydriven.pdf>.

GALVES, C. O enfraquecimento da concordância em português brasileiro. In: ROBERTS, I.; KATO, M. A. (Org.). *Português Brasileiro, uma viagem diacrônica*. Campinas: Editora da Unicamp, 1993. p. 387-408.

GALVES, C. *Ensaio sobre as gramáticas do português*. Campinas: Editora da Unicamp, 2001.

GONÇALVES, A.; SANTOS, A. L.; DUARTE, I. (Pseudo-)Inflected infinitives and Control as Agree. In: LAHOUSSE, K.; MARZO, S. *Romance Languages and Linguistic Theory: Selected papers from 'Going Romance'* Leuven. Amsterdam: John Benjamins, 2014. p. 161-180. <https://doi.org/10.1075/rllt.6.08gon>

GRANO, T. A. *Control and restructuring at the syntax-semantics interface*. Dissertation (Doctoral) – University of Chicago, Chicago, 2012.

GRANO, T. A. *Control and restructuring*. Oxford: Oxford University Press, 2015. <https://doi.org/10.1093/acprof:oso/9780198703921.001.0001>

GUY, G. R. Linguistic variation in Brazilian Portuguese: Aspects of the phonology, syntax, and language history. 1981. Dissertation (Doctoral) - University of Pennsylvania, Philadelphia, 1981.

HORNSTEIN, N. Movement and Control. *Linguistic Inquiry*, MIT Press, v. 30, p. 69-96, 1999. <https://doi.org/10.1162/002438999553968>

ILARI, R.; BASSO, R. *O português da gente: a língua que estudamos, a língua que falamos*. São Paulo: Ed. Contexto, 2006.

KATO, M. A. Strong and weak pronominals and the null subject parameter. *Probus*, Berlin, v. 11, n. 1, p. 1-38, 1999. <https://doi.org/10.1515/prbs.1999.11.1.1>

KATO, M.; NEGRÃO, E. V. (Org.). *Brazilian Portuguese and the null subject parameter*. Frankfurt: Vervuert-Iberoamericana, 2000.

KROCH, A. Syntactic Change. In: BALTIN, M.; COLLINS, C. (Org.). *The Handbook of Contemporary Syntactic Theory*. Oxford, Mass.: Blackwell Publishers Inc., 2001. p. 699-729. <https://doi.org/10.1002/9780470756416.ch22>

LANDAU, I. *Elements of Control: Structure and meaning in infinitival constructions*. Dordrecht: Kluwer, 2000. <https://doi.org/10.1007/978-94-011-3943-4>

LANDAU, I. Severing the Distribution of PRO from Case. *Syntax*, Wiley Online Library, v. 9, p. 153-170, 2006.

LANDAU, I. Two Routes of Control: Evidence from Case Transmission in Russian. *Natural Language and Linguistic Theory*, Springer Link, v. 26, p. 877-924, 2008.

LANDAU, I. *Control in generative grammar: a research companion*. Cambridge: Cambridge University Press, 2013. <https://doi.org/10.1017/CBO9781139061858>

LANDAU, I. *A two-tiered theory of Control*. Cambridge, MA: The MIT Press, 2015. <https://doi.org/10.7551/mitpress/9780262028851.001.0001>

LEMLE, M. *Análise sintática*. São Paulo: Ática, 1984.

LIGHTFOOT, D. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press, 1991.

LUCCHESI, D. As duas grandes vertentes da história sociolinguística do Brasil. *DELTA*, São Paulo, PUC-SP, v. 17, p. 97-130, 2001.

LUCCHESI, D. *Sistema, mudança e linguagem: um percurso na história da linguística moderna*. São Paulo: Parábola, 2004.

LUCCHESI, D. A concordância verbal e a polarização sociolinguística do Brasil. UFBA, 2012. Unpublished paper.

LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. *O Português Afro-Brasileiro*. Salvador: EDUFBA, 2009. <https://doi.org/10.7476/9788523208752>

MADEIRA, A. M. On the Portuguese inflected infinitive. In: XX INCONTRO DI GRAMMATICA GENERATIVA, XX., 1994. Pádua. *Teoria del Linguaggio e Analisi Lingüística*. Pádua: Unipress, 1994. p. 289-314.

MATTOS E SILVA, R. V. *Uma sócio-história do português brasileiro*. São Paulo: Parábola Editora, 2004.

MAURER Jr., T. M. *O infinitivo flexionado português*. São Paulo: Ed. Nacional; EdUSP, 1968.

McELREE, B.; BEVER, T. G. The psychological reality of linguistically defined gaps. *Journal of Psycholinguistic Research*, Springer Link, v. 18, p. 21-35, 1989.

McFADDEN, T.; SUNDARESAN, S. Finiteness in South Asian languages: an introduction. *Natural Language and Linguistic Theory*, Springer Link, v. 32, p. 1-28, 2014.

MILLER, D. G. *Nonfinite Structures in Theory and Change*. Oxford: Oxford University Press, 2002.

MODESTO, M. On the identification of null arguments. Dissertation (Doctoral) - University of Southern California (USC), Los Angeles, 2000a.

MODESTO, M. Null subjects without “rich” agreement. In: KATO, M. A.; NEGRÃO, E. V. (Org.). *Brazilian Portuguese and the null subject parameter*. Frankfurt: Vervuert- Iberoamericana, 2000b. p. 147-174.

MODESTO, M. Topic prominence and null subjects. In: BIBERAUER, T. (Org.). *The limits of syntactic variation*. Amsterdam: John Benjamins Publishing, 2008. p. 375-410. <https://doi.org/10.1075/la.132.17mod>.

MODESTO, M. What Brazilian Portuguese says about control: remarks on Boeckx & Hornstein. *Syntax*, Wiley Online Library, v. 13, n. 1, p. 78-96, 2010. <https://doi.org/10.1111/j.1467-9612.2009.00136.x>

MODESTO, M. Finite control: Where movement goes wrong in Brazilian Portuguese. *Journal of Portuguese Linguistics*, Lisboa, Ubiquity Press, v. 10, n. 1, p. 3-30. 2011.

MODESTO, M. Inflected infinitives in Brazilian Portuguese and the theory of control. In: GONÇALVES, A. (Org.). *Complement Clauses in Portuguese: Syntax and Acquisition*. Amsterdam: John Benjamins. 2016a. In press.

MODESTO, M. Inflected infinitives and restructuring in Brazilian Portuguese. In: KATO, M. A.; ORDÓÑEZ F. (Org.). *Morphosyntax of Spanish and Portuguese in Latin America*. Oxford: Oxford University Press, 2016b. p. 157-176. <https://doi.org/10.1093/acprof:oso/9780190465889.003.0007>.

MOREIRA DA SILVA, S. Etudes sur la symétrie et l'asymétrie SUJET/OBJET dans le Portugais du Brésil. 1984. Dissertation (Doctoral) - Université de Paris VIII, Paris, 1984.

NARO, A.; SCHERRE, M. M. P. *Origens do português brasileiro*. São Paulo: Parábola Editorial, 2007.

NEGRÃO, E. V. *Anaphora in Brazilian Portuguese complement structures*. Dissertation (Doctoral) - University of Wisconsin, Madison, 1986.

NEGRÃO, E. V. O português brasileiro: uma língua voltada para o discurso. 1999. Tese (Livre-docência) - Universidade de São Paulo, São Paulo, 1999.

NEGRÃO, E. V.; VIOTTI, E. Brazilian Portuguese as a Discourse-Oriented Language. In: KATO, M. A.; NEGRÃO, E. V. (Org.). *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt: Vervuert-Iberoamericana, 2000. p. 105-125.

NICOL, J.; SWINNEY, D. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, Springer Link, v. 18, p. 5-19, 1989.

NUNES, J. Inherent case as a licensing condition for A-movement: The case of hyper-raising constructions in Brazilian Portuguese. *Journal of Portuguese Linguistics*, Lisboa, Ubiquity Press, v. 7, n. 2, p. 83-108, 2008.

OSTERHOUT, L.; NICOL, J. The time-course of antecedent activation following empty subjects. 1988. Unpublished manuscript.

OUSHIRO, L. *Identidade na Pluralidade*. Tese (Doctorado) - Universidade de São Paulo, São Paulo, 2015.

PIRES, A.; ROTHMAN, J. Building bridges: Experimental L1 acquisition meets diachronic linguistics. In: GUIJARRO-FUENTES, P.; DOMÍNGUEZ, L. (Org.). *New directions in language acquisition: Romance languages in the generative perspective*. Newcastle upon Tyne, UK: Cambridge Scholars Publishing, 2010. p. 357-385.

PONTES, E. *O tópico no português do Brasil*. Campinas: Pontes, 1987.

POSSENTI, S. *Por que (não) ensinar gramática na escola*. Campinas: Mercado de Letras, 1996.

QUICOLI, A. C. Inflection and parametric variation: Portuguese vs. Spanish. In: FREIDIN, R. (Org.). *Current issues in comparative grammar*. Amsterdam: Springer, 1996. p. 46-80. https://doi.org/10.1007/978-94-009-0135-3_2

RAPOSO, E. Case theory and Infl-to-Comp: The inflected infinitive in European Portuguese. *Linguistic inquiry*, MIT Press, v. 18, n. 1, p. 85-109, 1987.

RAPOSO, E. Prepositional infinitival constructions in European Portuguese. In: JAEGGLI, O.; SAFIR, K. (Org.). *The null subject parameter*. Amsterdam: Springer, 1989. p. 277-305. https://doi.org/10.1007/978-94-009-2540-3_10

ROBERTS, I.; KATO, M. A. (Org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. da Unicamp, 1993.

RODRIGUES, C. Impoverished morphology and A-movement out of Case domains. Dissertation (Doctoral) - University of Maryland, College Park, 2004.

RODRIGUES, C. Agreement and Flotation in Partial and Inverse Partial Control Configurations. In: DUBINSKY, S.; DAVIES, W. (Org.). *New Horizons in the Analysis of Control and Raising*. Dordrecht: Springer, 2007. p. 213-229. https://doi.org/10.1007/978-1-4020-6176-9_9

RODRIGUES, C.; HORNSTEIN, N. Epicene Agreement and Inflected Infinitives When the Data Is “Under Control”: A Reply to Modesto (2010). *Syntax*, Wiley Online Library, v. 16, p. 292-309, 2013.

ROSENBAUM, P. *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: The MIT Press. 1967.

ROTHMAN, J.; DUARTE, I.; PIRES, A.; SANTOS, A. L. How early after all?: Inflected infinitives in European and Brazilian Portuguese L1 production. In: STAVRAKAKI, S.; LALIOTI, M.; KONSTATINOPOULOU, P. (Org.). *Advances in Language Acquisition*. Newcastle upon Tyne: Cambridge Scholars Publishing, 2013.

SCIDA, E. *The inflected infinitive in Romance Languages*. New York: Routledge Taylor, 2004.

SCHERRE, M. P.; NARO, A. J. Mudança sem mudança: a concordância de número no português brasileiro. *SCRIPTA*, Belo Horizonte, PUC Minas, v. 18, p. 107-129, 2006.

SHEEHAN, M. Partial Control in Romance Languages: the covert comitative analysis. In: LAHOUSSE, K.; MARZO, S. (Org.). *Romance Languages and Linguistic Theory*. Amsterdam: John Benjamins, 2014. p. 181-198. <https://doi.org/10.1075/rllt.6.09she>

SHEEHAN, M. Partial pro-drop. In: A. GONÇALVES (Org.). *Complement Clauses in Portuguese: Syntax and Acquisition*. Amsterdam: John Benjamins, 2016. In press.

SIGURÐSSON, H. A. The Case of PRO. *Natural Language and Linguistic Theory*, Springer Link, v.26, p. 403-450, 2008.

SILVA NETO, S. da. *Introdução ao estudo da língua portuguesa no Brasil*. Rio de Janeiro: Ed. Presença, 1950.

SITARIDOU, I. The synchrony and diachrony of Romance Infinitives with nominative subjects. Dissertation (Doctoral) – University of Manchester, Manchester, UK, 2002.

SITARIDOU, I. Romance infinitives with subjects, subjunctive obviation and control theory. In: EUGUREN, L.; FERNÁNDEZ-SORIANO, O. (Org.). *Coreference, modality, and Focus*. Amsterdam: John Benjamins, 2007. p. 191-219. <https://doi.org/10.1075/la.111.10sit>

SUNDARESAN, S.; McFADDEN, T. Subject distribution in Tamil and other languages: selection vs. case. *Journal of South Asian Languages*, De Gruyter, v. 2, p. 5-34, 2009.

SUNDARESAN, S. Making sense of silence: finiteness and the (OC) pro vs. pro distinction. *Natural Language and Linguistic Theory*, Springer Link, v. 32, p. 59-85, 2014.

SZABOLCSI, A. Overt nominative subjects in infinitival complements: Data, diagnostics, and preliminary analyses. In: IRWIN, P.; MALDONADO, V. V. R. (Org.). *Papers in syntax: NYU Working Papers in Linguistics*. New York: New York University, 2009. v. 2, p. 1-55.

TARALLO, F. On the alleged Creole origin of Brazilian Portuguese: untargeted syntactic change. *Cadernos de Estudos Lingüísticos*, Campinas, Unicamp, v. 15, p. 137-161, 1988.

TARALLO, F. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. In: ROBERTS I.; KATO, M. A. (Org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Editora da Unicamp. 1993. p. 69-105.

TROTZKE, A.; BADER, M.; FRAZIER, L. Third Factors and the Performance Interface in Language Design. *Biolinguistics*, Cyprus, University of Cyprus v.7, p. 1-34, 2013. Available at: <<http://www.biolinguistics.eu>>.

WURMBRAND, S. Restructuring cross-linguistically. In: NORTH EASTERN LINGUISTICS SOCIETY ANNUAL MEETING, 45., 2015, Amherst. BUI, T.; ÖZYILDIZ, D. (Org.). *Proceedings...* Amherst: University of Massachusetts, GLSA, 2015. p. 227-240.

On the asymmetry between subject and object relative clauses in discourse context¹

Assimetria no custo de processamento de relativas de sujeito e de objeto em contextos discursivos

Renê Forster

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
rene.forster@gmail.com

Letícia Maria Sicuro Corrêa

Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
lscorrea@puc-rio.br

Abstract: This paper investigates the possibility of an effect of contextual information during the processing of sentences containing subject relative clauses (SRCs) and object relative clauses (ORCs) in Brazilian Portuguese. The predictions from one-stage models and from syntax-oriented approaches to sentence processing are outlined. An eye-tracking experiment is reported in which SRCs and ORC were presented when preceded by narrative contexts that could either favor a subject or an object relative clause analysis. The results suggest that ORCs are harder

¹ This paper is an extended version of the work *On the unbalance between subject and object relative clauses in discourse context* presented in the XXIX CUNY Conference on human sentence processing, Florida, March 3-5, 2016. This research was supported by a post-doctoral scholarship for the first author (PDJ-CNPq) and was developed in connection with the project “*Processamento e aquisição da linguagem sob ótica minimalista: extensão e comparação de modelos*” (PQ-CNPq) of the second author.

to process when compared to SRCs, no matter what discourse contexts they are inserted in. The contextual effect obtained here can be ascribed to a pre-syntactic priming, ie. a priming effect which arises during lexical access. The possibility of pre- and post-syntactic contextual effects in the processing of RCs is discussed.

Keywords: relative clauses; sentence-processing; eye-tracking.

Resumo: Este trabalho investiga, no Português Brasileiro, a possibilidade de um efeito de informação contextual no processamento de sentenças contendo orações relativas de sujeito (RS) e de objeto (RO), considerando os possíveis impactos dessa integração para a assimetria de custo entre RSs e ROs. Previsões de modelos de base estrutural e de modelos interativos são comparadas. Um experimento de rastreamento ocular é relatado em que relativas de sujeito e de objeto foram apresentadas em contextos narrativos que favoreceriam o processamento, ou de uma relativa de sujeito, ou de uma relativa de objeto. Os resultados sugerem que ROs têm maior custo quando comparadas a relativas de sujeito, independentemente do contexto discursivo no qual estão inseridas. O efeito de informação contextual obtido pode ser atribuído a um efeito de *priming* pré-sintático, ou seja, um efeito de *priming* durante o acesso lexical. Discute-se a possibilidade de efeitos contextuais pré e pós-sintáticos no processamento de orações relativas.

Palavras-chave: orações relativas; compreensão da linguagem; rastreamento ocular.

Recebido em: 8 de janeiro de 2017.

Aprovado em: 24 de abril de 2017.

1 Introduction

The asymmetry between subject relative clauses (SRCs) and object relative clauses (ORCs) has been widely reported, with higher processing costs for the latter (eg. WANNER; MARATSOS, 1978; KING; JUST, 1991; KING; KUTAS, 1995; GIBSON *et al.*, 2005). This asymmetry has been ascribed to the transient memory load between the head of the RC (the filler) and the position in which it is recovered (the gap in the RC), in order for the processing of relativization to be accomplished (cf. (1-2)). The greater the distance between the filler and the gap, the greater the transient memory load (WANNER; MARATSOS, 1978). The need to hold the head noun in working memory as soon as the relative clause is identified and to process a subject-verb relation would account for the greater transient memory load between the filler and the gap of object RCs (CORRÊA, 1986, 1995). In light of recent linguistic accounts (RIZZI, 2006, 2013), the intervention of the RC subject in ORCs would account for this asymmetry, and explain the particular difficulty of children and language impaired individuals in coping with the processing of these sentences (GRILLO, 2009; FRIEDMANN, BELLETTI; RIZZI, 2009). In any case, the structural properties of ORCs would account for the asymmetry.

(1) ... **the boy** that ___ called the girl... (SRC)



(2) ... **the boy** that the girl called ___.... (ORC)



It is a point of debate, however, the extent to which the processing cost of relative clauses can be strictly attributed to their structural properties. It has been observed, for instance, that nonstructural factors such as *animacy*² (MAK; VONK; SCHRIEFERS, 2002, 2006; TRAXLER; MORRIS; SEELY, 2002) and *referential accessibility*

² It is worth noting that if *animacy* is assumed to be a formal feature, animacy effects on RC processing can be accounted for by the intervention hypothesis, i.e., it can be predicted by a structural approach to language comprehension (COSTA *et al.*, 2012).

(WARREN; GIBSON, 2005) of the DPs involved in the RC can be reflected on processing costs of this type of sentences. Some results suggest further that the asymmetry between SRCs and ORCs can disappear in narrative contexts (YANG; MO; LOUWERSE, 2013).

Addressing this debate, a study using the eye-tracking methodology was conducted with restrictive ORCs (AUGUSTO; CORRÊA; FORSTER, 2012; FORSTER, 2013). An effect suggestive of the integration of previous information during the analysis of RCs was obtained, which may have some impact on the overall processing cost. In that study, there were two characters (two girls) and two background events for the processing of the test sentences. In each of the events, one of the girls was the patient of the action performed by two similar/different characters. Eye-fixations reveal that, when the previous actions were performed by different characters, the participants tended to anticipate the referent of the object of the RC, as soon as its subject (the actor of one of the previous actions) was processed, by looking immediately and longer to the correct patient (one the two girls). This effect is compatible with results suggesting the anticipation of the referent of linguistic expression in biasing contexts (ALTMANN; KAMIDE, 1999).

A similar effect was not, however, obtained in a temporarily ambiguous context, i.e. when the sentence introduced by the complementizer/relative pronoun “that” could be preferentially interpreted as a complement sentence (FORSTER, 2013; FORSTER; CORRÊA, 2016) (cf. (3)). It seems, therefore, that a possible effect of information integration is blocked by the priority of structural factors – the minimal attachment of the embedded sentence (or the immediate satisfaction of the syntactic/semantic requirements of the main clause verb).

(3) Mary told the girl that the painter kissed...

In Forster (2013), two explanations were proposed as a means of reconciling the on-line integration of previous information with an autonomous parser. One possibility would be that the anticipation of the patient of the event described by the RC would facilitate the retrieval of the head noun at the position of the gap. This retrieval would be a means of checking the prediction prompted by the identification of the subject

of the RC (as in a sort of shadowing task³, as suggested in CORRÊA *et al.*, 2012). Another possibility would be a post-syntactic *priming* effect (FORSTER, 2013). The background context would prime the actor-action-patient relationship presented by the RC. Once the subject of the RC was recognized, the recovery of the patient (the head noun) would be facilitated, which could, in principle, minimize possible feature interference effects. In both cases, in order for the priority of the syntactic analysis to be maintained, the immediate mapping of the subject of the RC onto a referent requires that DPs are processed as *phases* i.e. chunks of syntactically analyzed material that are transferred to the interfaces between (internal) language and processing systems (CHOMSKY, 2001) in a model of on-line computation in which sentences are analyzed from left to right (CORRÊA *et al.*, 2012).

The present study resumes this debate by considering the extent to which a supportive context can override the asymmetry between SRCs and ORCs. An eye-tracking experiment was conducted, which was based on Yang *et al.*'s study (2013). Our results indicate that the asymmetry between ORCs and SRCs is unaffected by the contextual information provided.

The aim of this paper is then two-fold: to verify the extent to which the manipulation of contextual information affects the asymmetry between SRCs and ORCs and to discuss the possibilities of pre and post-syntactic contextual effects in RC processing.

This paper is organized as follows. In the next section, predictions from one-stage and from syntax-oriented models are briefly outlined. In section 3, the study conducted by Yang *et al.* (2013) is presented, in which the asymmetry between subject and object RCs is claimed to be neutralized.

³ The shadowing technique in elicited production tasks requires the participants to repeat as quickly as possible orally presented linguistic material (isolated sentences, normal prose) as it is heard, as in simultaneous interpretation from one language to another (MARSLÉN-WILSON, 1973a, b). The latency between the auditory stimulus and the onset of the repetition provides an estimate of the speed of the perception and the analysis of the linguistic material. The mean latency in the shadowing of connected speech is 250-300ms and there is evidence of anticipation (the participant produces something that has not already been heard, based on the on-going lexical/syntactic information). This type of result has been taken an indication that semantic and contextual constraints affect syntactic processing (TRUESWELL; TANNENHAUS, 1994). Alternatively, processing the material as hearer and as speaker may occur in parallel and anticipation by the speaker may be independently checked by the parser.

It is argued that the interpretation of the results is not as clear as suggested, which motivates the experiment reported in section 4. In section 5, possible discourse context effects in the processing of sentences with RCs are considered, which are in tune with current developments of linguistic theory. The final remarks present some future directions of the current study.

2 Integrating discourse information: two accounts for language comprehension

It can be said that two main approaches to the integration of nonstructural information in sentence processing have emerged throughout the history of psycholinguistics: syntax-based approaches, giving rise to a two-stage model, and constraint-based approaches, giving rise to one stage-models. By positing an encapsulated stage for structural analysis, syntax-based accounts predict that structural factors, such as the nature of the long distance dependencies involved, are one of the main sources of processing load. According to some of those models, first pass analyses guided by syntactic information may lead to incorrect structural descriptions. In RC processing, parsing heuristics, such as the *active filler strategy* (CLIFTON; FRAZIER, 1989; FRAZIER; FLORES D'ARCAIS, 1989; STOWE, 1986) would create a subject-bias for relative clauses since the possibility of a subject gap would be ranked over the option of a lexically occupied subject position. This strategy leads to incorrect structural descriptions for ORCs. Reanalysis, then, would increase processing costs. Models in this tradition predict that the asymmetry between ORCs and SRCs tends to prevail irrespectively of contextual information (FRAZIER; FODOR, 1978). More recent linguistically oriented proposals ascribe this asymmetry to an overextension of a universal principle that constrains the sort of local relations that can be established in a sentence (*Relativized Minimality*) (RIZZI, 2006, 2013). Even though the relationship between the head of an RC and its (phonologically null) copy (in the position of the gap) is licensed in human languages, there are conditions in which intervention occurs, thereby impairing the processing of ORCs (GRILLO, 2008, 2009; FRIEDMANN; BELLETI; RIZZI, 2009; FRIEDMANN; COSTA, 2010). The effect of intervention varies as a function of the syntactic properties of the intervening element (the subject of ORCs) – the greater the similarity between its syntactic properties and those of the

relativized element, more likely it is that intervention occurs (particularly in children and language impaired individuals who may not have access to the distinguishing WH-feature of the RC-object). In any case, insofar intervention is characterized as an interface phenomenon, the possibility of this effect being sensitive to contextual factors is an empirical question.

Approaches that rely on a general-purpose processor, with no exclusive stage for syntactic computation, claim that non structural information come into play since the very first steps of structural analysis (eg., MACDONALD; PEARLMUTTER; SEIDENBERG, 1994; TRUESWELL; TANENHAUS, 1994; STEVENSON, 1994; SPIVEY; TANENHAUS 1998; Cf. MCRAE; MATSUKI, 2012). According to one-stage models of sentence processing in constraint-based theories, initial structural descriptions can be assigned based on interpretive information, such as discourse information. Background information would then constrain the structure assigned to the currently processed material. These accounts predict that the processing cost associated with structural factors can be overridden by properties of the narrative context in which a sentence is inserted.

3 Reading relative clauses in context: Yang *et al.*'s study

The results obtained by Yang *et al.* (2013) in an eye-tracking experiment, are viewed as providing support for this latter account of language comprehension. In their study, the referent of the subject-DP of relative clauses was repeatedly presented in subject positions along a short narrative. The idea behind the manipulation was to create supportive contexts for either SRCs or ORCs. It seems that the rationale was the following: when reading the relative clause in a sentence like (5) after narratives like (7), readers do not fall in the trap of expecting a subject gap in the RC, since a subject function for *babysitter* (the proper subject of the ORC) was consistently emphasized along the story previously read. The same would hold true for sentences like (4) when read after stories like (6).

For gaze duration measures, a general facilitatory effect of contextual information was reported such that null contexts (test sentences in isolation) produced longer fixation times (see Fig. 1). This effect can be explained in terms of a purely lexical (pre-syntactic) priming mechanism that benefits from recurrent activation of a DP in previous

discourse in line with previous observations (GRODNER *et al.*, 2005). That is, the fact that the referents of the subject and the object of the RC are reiterated along the narrative passage can make lexical recognition and the access to the features of the lexical items easier in the analysis of the test sentences. It seems, then, that the effect of gaze duration does not say much about the attribution of structural descriptions.

In Yang *et al.*'s results the supportive contexts favoring SRCs led to longer reading times for ORCs when compared to SRCs, as expected. The typical asymmetry between ORCs and SRCs disappeared when the latter were read after supportive contexts intended to favor ORCs (see Fig. 1).

(4) SRC: *The child that chased the babysitter squealed with delight at the game.*

(5) ORC: *The child that the babysitter chased squealed with delight at the game.*

(6) Supportive context for SRCs: *The child with a pony tail was playing on the playground with the babysitter. She looked pretty good in the pink sweater, and her white sport shoes were very cute. When the child laughed, her voice sounded sweet. On this sunny spring day, the breeze gently blew her brown hair. The child was busy picking up wildflowers, chasing the birds, and playing a game of tag.*

(7) Supportive context for ORCs: *The babysitter with a pony tail was playing on the playground with the child. She looked pretty good in the pink sweater, and her white sport shoes were very cute. When the babysitter laughed, her voice sounded sweet. On this sunny spring day, the breeze gently blew the brown hair. The babysitter was busy picking up wildflowers, chasing the birds, and playing a game of tag.*

A closer look at Yang *et al.*'s data (see Fig. 1), however, calls into question the claim that the contextual information provided overrides the asymmetry between SRC and ORCS. It can be observed that ORCs are not read faster in contexts intended to favor ORCs when compared to contexts intended to favor SRCs. It seems, instead, that the processing

of SRCs becomes relatively more difficult in contexts favoring ORCs (cf. Figure 1). It is not clear, then, whether the conclusion “the object relative clause... becomes easier to read than the subject relative clause” (YANG *et al.*, 2013, p. 236) can be sustained.

FIGURE 1 – Results from Yang *et al.* (2013)



Gaze duration (left) and total fixation times (right) for relative clauses per type of narrative contexts

It is important to note that the ORC favoring condition is the only one, according to Yang *et al.*'s data, in which the typical asymmetry between SRCs and ORCs fails to be attested. According to the authors, this result is due to the consistent association of the referent of the RC subject-DP with the subject function. In the absence of the reiterated recovery of a given referent by a subject RC in the supportive context provided, as in Fedorenko *et al.* (2012), the asymmetry between ORCs and SRCs prevails. Fedorenko *et al.* (2012) tested relative clauses like (8) and (9) in a self-paced reading experiment. The typical asymmetry between ORCs and SRCs was observed with or without the presence of previous short narratives like (10).

(8) I heard that the reporter that attacked the senator admitted to making an error.

(9) I heard that the reporter that the senator attacked admitted to making an error

(10) At the press-conference, a senator and two reporters got into an argument. The senator attacked one of the reporters and then the other reporter attacked the senator.

An alternative explanation for the relative difficulty of SRCs in ORC favoring context and the absence of a facilitative effect in ORCs can be provided. Given the sentence (4) in context (7), the sudden reintroduction of the DP *the child* after several mentions of the concurrent DP *the babysitter* may explain the longer reading times for SRCs in the context condition. If this is the case, it can be argued that the parsing of the embedded sentence was not necessarily affected by the contextual information provided.

4 Experiment

4.1 Design and predictions

The experiment reported here was intended to make the contrast between supportive contexts clearer than in Yang *et al.*'s study in order to verify the extent to which a biasing context can override the asymmetry between subject and object RCs.

As in that study, the eye-tracking methodology was used in a reading comprehension task. Unlike that study, the present one focused only on the comparison between SRCs and ORCs in two different contexts. The null condition was excluded here. The sort of experimental material used in that study was nevertheless maintained. The supportive contexts were characterized in relation to the element that is consistently presented by a DP-subject in the narrative preceding the test sentences, that is, the referent of the subject of the RC. As such, Context A was intended to favor the processing of SRC and Context B the processing of ORCs, as in Yang *et al.*

The reading task consisted of the silent reading of a narrative in which the critical sentence was included. A comprehension question was to be answered based on this story. The reading task and the comprehension question were followed by a distractor rating task, in which the participants should evaluate how coherent or sound the story was, using a 1 (very bad) to 5 (very good) scale, insofar as the

experimental material, following the original study, might not sound natural in some cases. After the rating task, subjects had to press the space bar to move to the next stimulus.

The independent variables were (i) the type of *RC structure* (SRC or ORC) ((11) & (12)) and (ii) type of *Narrative context* (Context A and Context B) ((13) & (14)), in a 2x2 factorial design. Context A emphasized the subject of SRC and Context B, the subject of ORCs.

(11) SRC: O banqueiro que [Δ irritou o advogado] joga tênis todo sábado.
The banker that irritated the lawyer plays tennis every Saturday.

(12) ORC: O banqueiro que [o advogado irritou Δ] joga tênis todo sábado.
The banker that the lawyer irritated plays tennis every Saturday.

(13) Context A - bias for SRC: O dia do banqueiro começou com uma rápida lida nos jornais para ver o que estava acontecendo no mundo. O telefone dele não parava de tocar. **Ele** atendia educadamente, mas [**ele**] estava esperando impacientemente a ligação do advogado. A ligação finalmente veio. **O banqueiro** falou no telefone por mais de uma hora. Apesar disso, eles não chegaram a um acordo. [**RELATIVA**]. Eles costumam se encontrar no clube de tênis. O banqueiro esperava encontrar com o advogado novamente neste sábado. (*The banker's day started with a quick reading of the newspapers to see what was going on in the world. His phone didn't stop ringing. He answered politely, but [he] was impatiently waiting for a call from the lawyer. The call finally came. The banker talked on the phone for more than one hour. Despite this, they didn't come to a deal. [RELATIVE CLAUSE]. They usually meet on the tennis club. The banker hoped to meet the lawyer again this Saturday.*)

(14) Context B – bias for ORC: O dia do advogado começou com uma rápida lida nos jornais para ver o que estava acontecendo no mundo. O telefone dele não parava de tocar. **Ele** atendeu educadamente, mas [**ele**] estava esperando impacientemente a ligação do banqueiro. A ligação finalmente veio. **O advogado** falou no telefone por mais de uma hora. Apesar disso, eles não chegaram a um acordo. [**RELATIVA**]. Eles

costumam se encontrar no clube de tênis. O advogado esperava não ter que cruzar com o banqueiro no clube. (The lawyer's day started with a quick reading of the newspapers to see what was going on in the world. His phone didn't stop ringing. **He** answered politely, but [**he**] was impatiently waiting for a call from the banker. The call finally came. **The lawyer** talked on the phone for more than one hour. Despite this, they didn't come to a deal. [**RELATIVE CLAUSE**]. They usually meet on the tennis club. The lawyer hoped to meet the banker again this Saturday.)

Three critical regions were delimited: the head noun region (head DP + relative pronoun); the RC region (the subject-verb-object structure following the relative pronoun, in which either the subject position or the object position present a gap); and the spillover region immediately after the RC, including the main verb.

(15) Head que [that] Subj. Verb Object (...) main clause verb
 1 2 3

1. Head noun region
2. RC region
3. Spill-over region

The relevant regions for comparing the present results with the results of Yang *et al.* are 2 and 3. Region 1 makes it possible to capture differences in reading times due to regressions. This region was also intended to provide means of verifying if changing the referent of the main clause subject could explain the longer reading times for SRC in the ORCs supportive context in Yang *et al.*'s study.

The dependent variables were (i) total fixation duration at the head noun region, (ii) total fixation duration at the RC region, and (iii) total fixation duration at the spillover region. By *total fixation duration*, we mean the sum of all fixations on a portion of the sentence comprehended between the moment that the eyes first fixate on this region to the moment the last fixation in that area is detected, including all regressions. Additionally, the number of correct responses to the comprehension questions was considered.

The following predictions were verified:

(a) If the effect of structure prevails, a main effect of RC structure is expected with longer total fixations in ORCs in the RC region and in the spillover region (the region which includes the main clause verb, where the complex subject is analyzed in relation with the verb).

(b) If the type of narrative context interferes in RC processing, in the RC region, an interaction between *RC structure* and *Context* should be obtained, in which the asymmetry between SRC and ORC is reverted or neutralized in Context B.

4.2 Method

PARTICIPANTS

The participants were 41 undergraduate students (30 female) from Pontifical Catholic University of Rio de Janeiro (PUC-Rio). Nine of them were excluded due to bad calibration, errors of the researcher or due to bad performance on comprehension questions. The participants signed a term of consent and volunteered to take part in the experiment. Their participation was exchanged for course credits. The mean age of the final sample was 22.15 years old.

MATERIAL

40 test narratives were constructed, 20 for each version (favoring SRC or favoring ORC), as close as possible to the original narratives in English. There were also 40 test-sentences containing relative clauses, 20 for each version (SRCs or ORCs). Besides the critical stimuli, an additional 10 fillers and 2 training narratives were created. Four lists were constructed in such a way that the participants could see each version of a relative clause only once. Materials were presented in computer screen in a monospaced font with each 1° of the visual field ranging approximately 4 characters.

APPARATUS

A 300 Hz eye-tracker (with 0.5° accuracy and 0.4° precision) was connected to a Windows desktop computer for monitoring participants while stimuli were presented in a 23" full-hd screen. The software *Tobii Studio* was used for presenting stimuli, collecting data and analyzing the results. Subjects' verbal responses to comprehension questions and rating tasks were recorded through a Samson USB microphone.

PROCEDURE

Subjects had to read text passages presented at the computer screen while their eye movements were recorded. The session was preceded by a calibration phase. In this phase, subjects were required to look at 9 points in the screen. Additional recalibration points were shown in case of bad calibration. The calibration phase was followed by a training phase, in which two narratives were presented. Then, if the participant demonstrated having understood the procedure, one of the four lists of randomly ordered items was presented. After each story, question and rating task, the space bar was to be pressed for the next stimulus. The comprehension questions and rating task were orally answered by the participant. There was a break in the middle of each session in which the participant was allowed to rest. The second half of the session restarted with another calibration phase, as soon as the participant signaled to be ready. The whole session lasted between 25 and 35 minutes.

4.3 Results

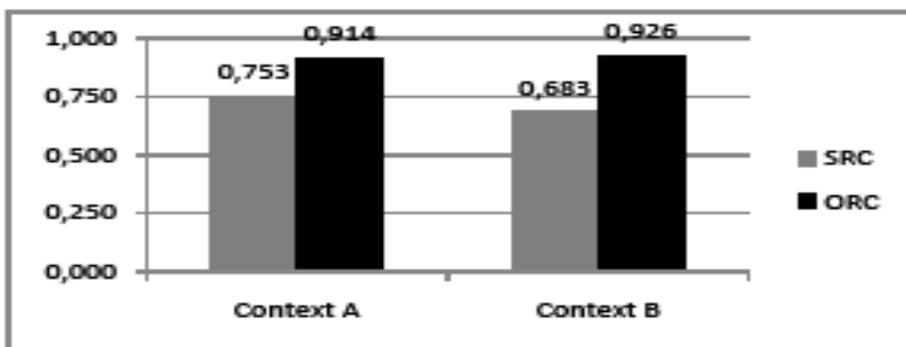
4.3.1 Comprehension question

There were 91.4% correct response to the comprehension questions. The comprehension task was, therefore, easily performed.

4.3.2 Reading time

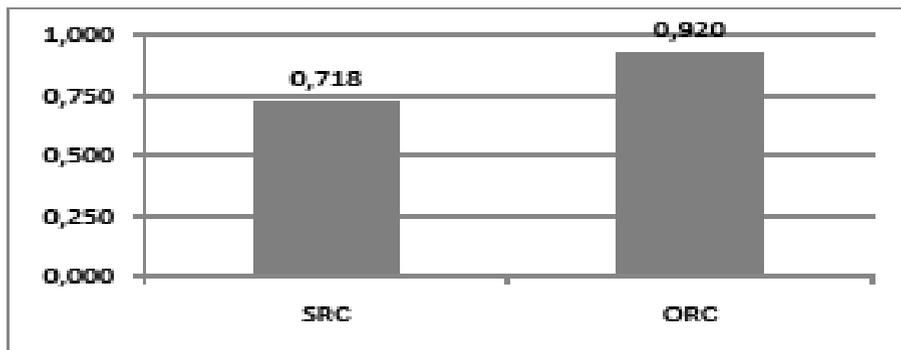
The data concerning the total fixation duration (ms) in each critical region were analyzed by means of ANOVAs. Graph 1 presents the mean total fixation duration in the head noun region in each condition. For this region, a main effect of *RC structure* was obtained ($F(1,31) = 13,6$ $p > .0009$), with longer reading times for ORCs ($M = 920$ ms) when compared to SRCs ($M = 718$), probably due to regressions during ORC reading (cf. Graph 2).

GRAPH 1– Mean total fixation duration at the head noun region per condition (ms)

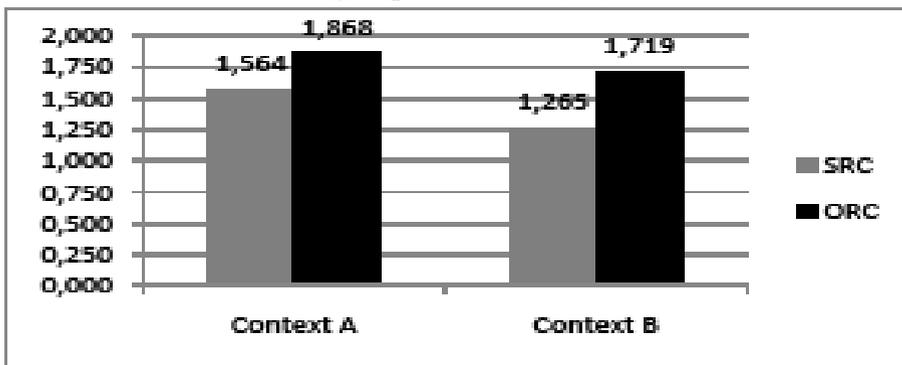


Context A: bias for subject RC | Context B: bias for object RC

GRAPH 2 – Mean total fixation duration at the head noun region per *Structure of RC* (ms)



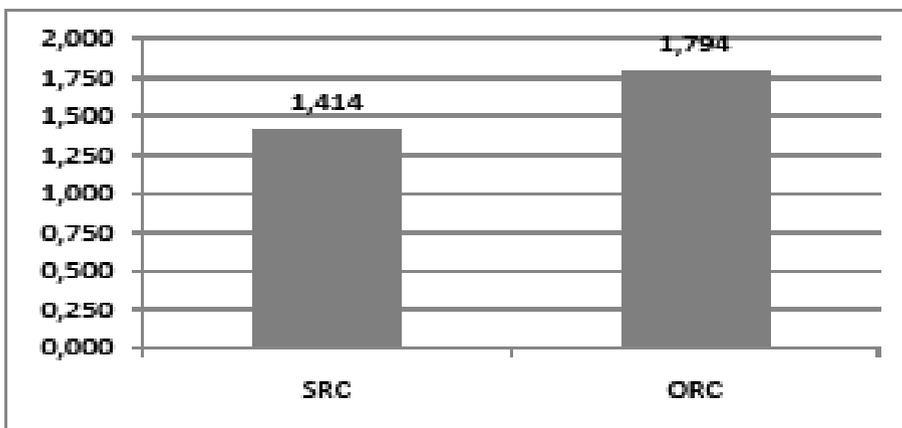
GRAPH 3 – Mean total fixation duration at the relative clause region per condition (ms)



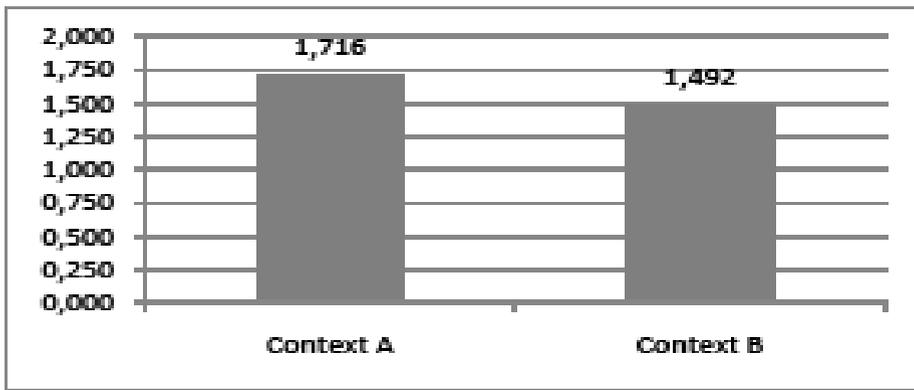
Context A: bias for subject RC | Context B: bias for object RC

For the Relative clause region, Graph 3 presents the mean total fixation duration per condition. A main effect of *RC structure* was obtained ($F_1(1,31) = 24,9$ $p < .00003$; $F_2(1,19) = 12,3$ $p < .003$) (cf. Graph 4). Longer reading times were associated with ORCs regardless of context. A main effect of *Narrative context* ($F_1(1,31) = 24,9$ $p < .001$; $F_2(1,19) = 6,47$ $p < .02$) was also obtained, with longer reading times in context A than in B (cf. Graph 5).

GRAPH 4 – Mean total fixation duration at the relative clause region per Structure of RC (ms)



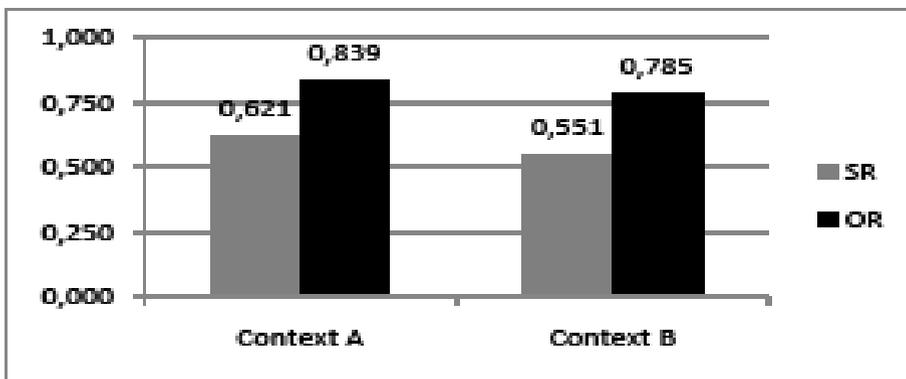
GRAPH 5 – Mean total fixation duration at the relative clause region per Context (ms)



Context A: bias for subject RC | Context B: bias for object RC

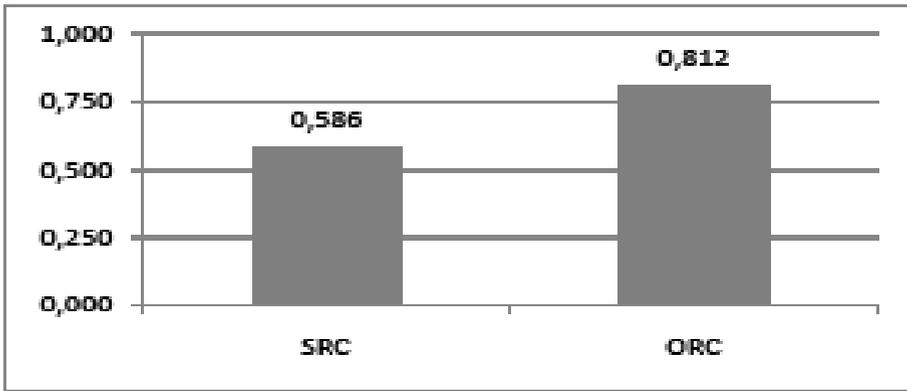
Graph 6 presents the mean fixation time for the *spill over* region. A main effect of *RC structure* ($F_1(1,31) = 35,1$ $p < .000003$; $F_2(1,19) = 15,9$ $p < .0008$) was obtained due to longer reading times for ORCs independently of *Narrative context* (cf. Graph 7). Neither a main effect of *narrative context* nor a significant interaction between variables was attested.

GRAPH 6 – Mean total fixation duration at the spill-over region per condition (ms)



Context A: bias for subject RC | Context B: bias for object RC

GRAPH 7 – Mean total fixation duration at the spill-over region per Structure of RC (ms)



4.4 Discussion

The direct comparison of SRC and ORC in contexts A and B provided a clearer picture of the role of a supportive context in which the subject of the RC is reiterated as the subject of the previous sentences. The results support prediction (a): ORCs had longer total fixation duration than SRCs regardless of *Context* in the three critical regions. No evidence was obtained for ORC processing being facilitated by supportive contexts in such a way as to revert or neutralize the asymmetry between SRCs and ORCs (prediction (b)). Unlike Yang *et al.*'s results, longer times were not obtained in SRC in Context 2 in the RC region and no effect of context could be captured in the head noun region that could suggest that altering the referent of the subject of the main clause could have impact in the processing of SRCs. The effect of RC structure in region 1 (the head noun region) is likely to be due to regressions, since the *total fixation duration* includes not only first-pass reading (ie. the first time the eye passes through a word), but also backward fixations detected on a critical region. This effect seems, then, to reflect processing costs of the RC region and, in particular, the higher cognitive demands of ORCs. A possible effect of the alternation of the subject of the main clause in the processing of SRCs in Yang *et al.*'s study was not captured here. It is possible that other eye-tracking measures (related to first-pass reading) or a self-paced reading task can provide further information about the effect of the reintroduction of a possibly unexpected DP.

Overall, these results suggest that the analysis of ORCs is more demanding than SRCs, no matter the particular context they are inserted in. Syntactic based models can explain these results as a possible feature intervention effect (GRILLO, 2008; FRIEDMANN; COSTA, 2010), which adds to the transient memory load and can lead to reanalysis (eg. CLIFTON; FRAZIER, 1989).

As for the effect of *Narrative context* in the RC region, it was not predicted. It does not seem to have affected the processing of the filler-gap relationship per se. This effect, unlike the effect of *RC structure*, was not maintained in the spill-over region (the matrix verb region), when the complex DP is processed as the subject of the verb. It is suggested then that the preceding context affected the processing of the critical sentence *at the level of lexical access only*. It seems that the repetition of the item that is overtly present in the RC in context B (*advogado*) (see the example (13) repeated below as (15)) has facilitated its recognition in both RC types ((16) and (17)), leading to the shorter reading times in the RCs in Context B. Facilitation is observed, then, when the DP phonetically visible inside the RC (and not the referent of the gap) is emphasized in prior discourse.

(15) Context A: **O dia do advogado** começou com uma rápida lida nos jornais para ver o que estava acontecendo no mundo. O telefone **dele** não parava de tocar. **Ele** atendeu educadamente, mas [**ele**] estava esperando impacientemente a ligação do banqueiro. A ligação finalmente veio. **O advogado** falou no telefone por mais de uma hora. Apesar disso, eles não chegaram a um acordo. [RELATIVA]. Eles costumam se encontrar no clube de tênis. O advogado esperava não ter que cruzar com o banqueiro no clube.

(16) SRC: O banqueiro que [Δ irritou o **advogado**] joga tênis todo sábado.

(17) ORC: O banqueiro que [o **advogado** irritou Δ] joga tênis todo sábado.

In sum, the present data support the view that the higher processing cost of ORCs is maintained regardless of context, as suggested by Fedorenko *et al.* (2012). Having the two types of supportive contexts envisaged in the Yang *et al.*'s study contrasted, in the absence of the null

context condition, provided a clearer assessment of the possibility of the reiterated mention of the subject of the RC in the subject of previous sentences alter the relative processing cost of sentences with RCs.

The more plausible explanation for the effect of context in Yang *et al.*'s study is that it was due to pre-syntactic lexical priming; that is, having the test sentences in a discourse context can facilitate lexical access in the RC. A spurious lexical effect along these lines may account for the main effect of context in the present study. It is also worth noting that in the narratives provided, the information conveyed by the RC was not necessarily crucial for the identification of a given referent in a set of referents of the same type. In the contexts exemplified in (6-7; 13-14), for instance, the restrictive RC singles out a single child and a single banker in the class of children and bankers, whereas in a context such as (10), the RC distinguishes a single reporter in a set of two reporters. The immediate search for the referent of the head/noun and the subject/object of the RC, as captured in Forster *et al.* (2011, 2013), which may be facilitated by a post-syntactic priming, is not crucially required in more descriptive RCs (as in 6-7). An effect of pre-syntactic lexical priming may therefore predominate, irrespective of the gap-filler relationship.

Given the prevalence of the effect of structure in the present results, potential contextual influence in RC processing can be due to either pre or post-syntactic effects. In the next section, the integration of RC information with the background context is characterized in the context of an on-line model of sentence computation, which incorporates an autonomous parser.

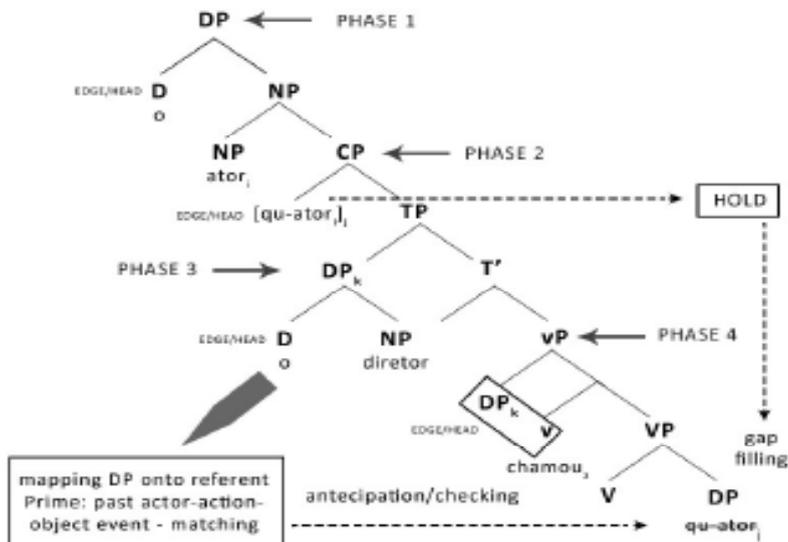
5 Pre and post-syntactic effects in the on-line processing of restrictive RCs

One of the main challenges for a syntactically oriented model of sentence processing is to account for the incremental mapping of DPs onto referents as the analysis of the sentence proceeds, which enables information from different sources to be integrated. Evidence for incrementality in sentence processing has motivated the proposal of single-stage constraint theories (eg. MACDONALD; PEARLMUTTER; SEIDENBERG, 1994; TRUESWELL; TANENHAUS, 1994; STEVENSON, 1994; SPIVEY; TANENHAUS 1998; Cf. MCRAE; MATSUKI, 2012). There is, nevertheless, compelling evidence for a syntax-first path analysis giving rise to small chunks of syntactically

analyzed material that can be semantically interpreted and mapped onto the referential world as the analysis of the sentence proceeds (from left to right) (VOS; FRIEDERICI, 2003).

It is our contention that the current developments of (Chomskyan) generative linguistics go in the direction of minimizing the differences that have been pulling apart constraint-based theories and dual-stage models of sentence processing (cf. CORRÊA *et al.*, 2012). Notice that universal principles that constrain possible syntactic relations in human languages are viewed as being imposed by performance systems (CHOMSKY, 1995, 2005; RIZZI, 2006, 2013) and all grammatical information is codified in the formal features of the lexical items. Formal features have interpretable and non-interpretable counterparts, which enable syntactic computation to be conducted modularly, while at the same time guaranteeing the interpretability of the product of the syntactic analysis. From a psycholinguistic point of view, having grammatical information represented as formal features of lexical items makes lexical access (which can be influenced by different factors) a precondition for syntactic processing. Moreover, computation (the merging of lexical items, giving rise to hierarchically structured objects) starts by the selection of items in an initial array or subarray. From a psycholinguistic point of view, the elements of these (sub) arrays are retrieved from the mental lexicon during on-line computation, in sentence production/comprehension. Derivation proceeds by phase (that is, as soon as there is a chunk of semantically interpretable syntactically analyzed material, the output of syntactic computation is transferred to the interfaces levels, i.e. it becomes accessible to processing systems (CHOMSKY, 2001). These developments of linguistic theory enable a model of on-line syntactic computation to be conceived of, which can account for incrementality in sentence processing while keeping an autonomous parser (CORRÊA; AUGUSTO, 2007, 2011). In particular, it has been argued that having DPs as *phases* in an on-line model of syntactic computation, in which phases are transferred from left to right (cf. AUGUSTO; CORRÊA; FORSTER, 2012) makes it possible to account for pre and post-syntactic effects in the comprehension of restrictive RCs.

FIGURE 2 – On line processing of restrictive ORC with post-syntactic priming



In Forster (2013), a procedural characterization of RC processing was proposed in which a post-syntactic effect of background context could be predicted. This account incorporates a processing unit equivalent to *phase*, and the HOLD model of RC comprehension in Corrêa (1986, 1995). According to that model, as soon as the relative pronoun is recognized, the closure of the current phrase (the DP) is postponed. The head noun is held active in working memory to be retrieved when a gap is found. In the present version, the holding of the head noun would be promoted by the recognition of the relative pronoun in the edge of a phase (cf. Fig 2).

This computation presupposes recognition and access to the lexical items presented in a sequence and analyzed from left to right. Pre-syntactic effects in sentence comprehension occur in the lexical recognition process. Once lexical items are recognized (via their phonological/written form), their formal and semantic features can be retrieved from the mental lexicon. This retrieval is amenable to the effect of frequency as well as to an effect of lexical priming when sentences are processed in discourse. The frequency of the items in language use, in the previous context, their salience in previous sentence and in the

discourse and the mere fact of having been presented can be reflected in processing speed (reading time, for instance).

Once recognized, the lexical items are merged into syntactic objects solely based on the grammatical information represented in their formal features (their semantic and phonological features are processed at the interface levels). As soon as uninterpretable features are eliminated (in the syntactic module), a chunk of information is sent to the interfaces. This syntactic object can be semantically interpreted and even mapped onto a referent/event. At the same time, its lexical features are kept active in working memory (edge features) so that further syntactic computation requiring it, such as gap filling operation, can occur. A DP mapped onto a referent (or an antecedent in discourse) may promote the retrieval of an event and its participants (introduced in the discourse by means of DPs). These DPs can then promote a post-syntactic priming effect. Keeping lexical features active in working memory makes them vulnerable to intervention effects once the analysis of the sentence proceeds (such as an immediate gap filler strategy).

Pre-syntactic effects in sentences processed in discourse contexts are expected to necessarily occur. Post-syntactic effects due to the mapping of DPs onto referents as the analysis of the sentence proceeds are likely to occur when the information in the RC is deeply integrated, that is, crucial for the identification of the referent of the head noun (as in CORRÊA *et al.*, 2012; FORSTER, 2013).

The extent to which an effect of feature intervention can be minimized in contexts such as (10) is still an open question. There is evidence that typical and language impaired children improve their ability to process object which-questions (vulnerable to intervention) when more than one possible referent to the head noun is available in the discourse context (CORRÊA *et al.*, 2012). In sum, pre-syntactic factors cannot affect the asymmetry between SRC and ORC, but the availability of lexical representations for syntactic computation to start. As for a post-syntactic effect, it is likely to be restricted to ORCs since the subject of a SRC can be immediately filled in as the verb is recognized. This is, nevertheless, an empirical question to be explored in the future.

6. Final remarks

The present results suggest that the differential demands of SRC and ORC are maintained in narrative contexts, at least in contexts in which the referent of the subject of the RC is consistently presented in subject position in the preceding discourse. The higher transient memory load between the head noun and the gap in ORCs (WANNER; MARATSOS, 1978; CORREA, 1995) possibly due to feature intervention (GRILLO, 2009) is an account that can be sustained based on the present results. Previous results suggesting the neutralization of this asymmetry allow alternative interpretations. Testing the impact of the main clause subject of a sentence with a center-embedded SRC in contexts favoring ORCs, by means of a self-paced reading task, is a follow-up planned to evaluate a possible alternative explanation for the asymmetry between SRC and ORCs observed in Yang's *et al.* results.

The main effect of contextual information obtained in the present study seems to stem from a pre-syntactic priming effect, which is likely to affect lexical access during discourse processing. Possible post-syntactic integration of restrictive RCs with prior information require, nevertheless, further investigation. A contrast between restrictive RCs that enables the identification of a single referent in a previously introduced set of individuals of the same type with restrictive RCs that describe a distinguishing property of a previously introduced referent can provide means of verifying the extent to which integrating the subject of an ORC with a past event affects processing costs.

References

AUGUSTO, M. R. A.; CORRÊA, L. M. S.; FORSTER, R. An argument for DPs as phases in an integrated model of on-line computation: the immediate mapping of complex DPs with relative clauses. *Revista Virtual de Estudos da Linguagem – ReVEL*, [s.l.], v. 10, n. 6, p. 7-26, 2012.

ALTMANN, G.; KAMIDE, Y. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, Elsevier, v. 73, p. 247-264, 1999. [https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)

CHOMSKY, Noam. *The minimalist program*. Cambridge, MA: MIT Press, 1995.

CHOMSKY, Noam. Derivation by Phase. In: KENSTOWICZ, M. (Ed.). *Ken Hale: a Life in Language*. Cambridge: MIT Press, 2001. p. 1-52.

CHOMSKY, Noam. Three factors in language design. *Linguistic Inquiry*, MIT Press, v. 36, n. 1, p. 1-22, 2005. <https://doi.org/10.1162/0024389052993655>

CLIFTON JR, C.; FRAZIER, L. Comprehending sentences with long-distance dependencies. In: CARLSON, G. N.; TANENHAUS, M. K. (Ed.). *Linguistic structure in language processing*. Netherlands: Springer, 1989. https://doi.org/10.1007/978-94-009-2729-2_8

CORRÊA, L. M. S. On the comprehension of Relative Clauses: a developmental study with reference to Portuguese. 1986. Thesis (PhD) – University of London, 1986.

CORRÊA, L. M. S. The Relative Difficulty of Children's Comprehension of Relative Clauses: A Procedural Account. In: NELSON, K.; RÈGER, Z. (Org.). *Children's Language*. Hillsdale, N. Jersey: Lawrence Erlbaum Assoc., 1995. v. 8, p. 225-244.

CORRÊA, L. M. S.; AUGUSTO, M. R. A. Computação linguística no processamento on-line: soluções formais para a incorporação de uma derivação minimalista em modelos de processamento. *Cadernos de Estudos Linguísticos*, Campinas, Unicamp, v. 49, p. 167-183, 2007.

CORRÊA, L. M. S.; AUGUSTO, M. R. A. Possible loci of SLI from a both linguistic and psycholinguistic perspective. *Lingua* (Haarlem. Print), Elsevier, v. 121, p. 476-486, 2011.

CORRÊA, L. M. S.; AUGUSTO, M. R. A.; LONGCHAMPS, J. R.; FORSTER, R. Referência Anafórica com relativas restritivas de objeto: custo relativizado na interface gramática-pragmática. *Revista Lingüística*, Rio de Janeiro, UFRJ, v. 8, n. 2, p. 173-196, 2012.

COSTA, João; GRILLO, Nino; LOBO, Maria. Minimality beyond lexical restrictions: processing and acquisition of free wh-dependencies in European Portuguese. *Revue Roumaine de Linguistique*, Institutul de Lingvistica al Academiei, v. 57, n. 2, p. 143-60, 2012.

FEDORENKO, E.; PIANTADOSI, S.; GIBSON, E. Processing relative clauses in supportive contexts. *Cognitive Science*, Wiley Online Library, v. 36, p. 471-497, 2012.

FORSTER, R. Aspectos do processamento de orações relativas: antecipação de referentes e integração de informação contextual. 2013. Tese (Doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013.

FORSTER, R.; CORRÊA, L. M. S. Rastreando o processamento de relativas de objeto: antecipação e integração de informação contextual na resolução de ambiguidades temporárias. *Gragoatá*, Niterói, UFF, v. 40, p. 133-156, 2016.

FRAZIER, L.; FLORES D'ARCAIS, F. G. Filler-Driven Parsing: A Study of Gap-Filling in Dutch. *Journal of Memory and Language*, Elsevier, v. 28, p. 331-344, 1989. [https://doi.org/10.1016/0749-596X\(89\)90037-5](https://doi.org/10.1016/0749-596X(89)90037-5)

FRAZIER, L.; FODOR, J. D. The Sausage machine: A new two-stage parsing model. *Cognition*, Elsevier, v.6, n. 4, p. 291-235, Dec. 1978. [https://doi.org/10.1016/0010-0277\(78\)90002-1](https://doi.org/10.1016/0010-0277(78)90002-1)

FRIEDMANN, N.; BELLETTI, A.; RIZZI, L. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua*, Elsevier, v. 119, n. 1, p. 67-88, 2009. <https://doi.org/10.1016/j.lingua.2008.09.002>

FRIEDMANN, N.; COSTA, J. The child heard a coordinated sentence and wondered: On children's difficulty in understanding coordination and relative clauses with crossing dependencies. *Lingua*, Elsevier, v. 120, n. 6, p. 1502-1515, 2010. <https://doi.org/10.1016/j.lingua.2009.10.006>

GIBSON, E.; DESMET, T.; GRODNER, D.; WATSON, D.; KO, K. Reading Relative Clauses in English. *Cognitive Linguistics*, De Gruyter, v. 16, p. 313-53, 2005.

GORDON, P. C, R. HENDRICK, M. JOHNSON. Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, Elsevier, v. 51, p. 97-114, 2004. <https://doi.org/10.1016/j.jml.2004.02.003>

GRILLO, N. Generalized minimality: syntactic underspecification in Broca's aphasia. 2008. Dissertation (PhD.) – University of Utrecht, 2008. <https://doi.org/10.1016/j.lingua.2008.04.003>

GRILLO, N. Generalized Minimality: Feature impoverishment and comprehension deficits in agrammatism. *Lingua*, Elsevier, v. 119, n. 10, p. 1426-1443, 2009.

GRODNER, D.; GIBSON, E.; WATSON, D. The influence of contextual contrast on syntactic processing: Evidence for strong-interaction in sentence comprehension. *Cognition*, Elsevier, v. 95, p. 275-296, 2005. <https://doi.org/10.1016/j.cognition.2004.01.007>

KING, J.; JUST, M. Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language*, Elsevier, v. 30, p. 580-602, 1991. [https://doi.org/10.1016/0749-596X\(91\)90027-H](https://doi.org/10.1016/0749-596X(91)90027-H)

KING, J.; W.; KUTAS, M. Who Did What and When? Using Word- and Clause-Level ERPs to Monitor Working Memory Usage in Reading. *Journal of Cognitive Neuroscience*, MIT Press, v. 7, p. 376-95, 1995. <https://doi.org/10.1162/jocn.1995.7.3.376>

MACDONALD, M. C.; PEARLMUTTER, N. J.; SEIDENBERG, M. S. The lexical nature of syntactic ambiguity resolution. *Psychological Review*, Elsevier, v. 101, n. 4, p. 676-703, Oct. 1994. <https://doi.org/10.1037/0033-295X.101.4.676>

MAK, W. M.; VONK, W.; SCHRIEFERS, H. The Influence of Animacy on Relative Clause Processing. *Journal of Memory and Language*, Elsevier, v. 47, p. 50-68, 2002. <https://doi.org/10.1006/jmla.2001.2837>

MAK, W. M.; VONK, W.; SCHRIEFERS, H. Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, Elsevier, v. 54, p. 466-490, 2006. <https://doi.org/10.1016/j.jml.2006.01.001>

MARSLEN-WILSON, W. D. Linguistic structure and speech shadowing at very short latencies. *Nature*, Springer Nature, v. 244, p. 522-523, 1973. <https://doi.org/10.1038/244522a0>

MARSLEN-WILSON, W. D. Speech shadowing and speech comprehension. *Speech Communication*, Elsevier, n. 4, n.1-3, p. 55-73, 1985. [https://doi.org/10.1016/0167-6393\(85\)90036-6](https://doi.org/10.1016/0167-6393(85)90036-6)

MCRAE, K.; MATSUKI, K. Constraint-based Models of Sentence Processing. In: VAN GOMPEL, R. (Ed.). *Sentence Processing*. New York: Psychology Press, 2012.

RIZZI, L. *Relativized Minimality*. Cambridge Mass.: MIT Press, 1990.

RIZZI, L. On the form of chains: Criterial positions and ECP effects. *Current Studies in Linguistics Series*, MIT Press, v. 42, p. 97, 2006.

RIZZI, L. Locality. *Lingua*, Elsevier, v. 130, p. 169-186, 2013.

SPIVEY, M. J.; TANENHAUS, M. K. Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory and Cognition*, Washington, v. 24, p. 1521-1543, 1998.

STEVENSON, S. Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, Springer Link, v. 23, p. 295-322, 1994.

STOWE, L. A. Parsing WH-constructions: Evidence for online gap location. *Language and Cognitive Processes*, Taylor & Francis Online, n.1, p. 227-245, 1986.

TRAXLER, M. J.; MORRIS, R. K.; SEELY, R. E. Processing subject and object relative clauses: Evidence from eye movement. *Journal of Memory and Language*, Elsevier, v. 47, p. 69-90, 2002. <https://doi.org/10.1006/jmla.2001.2836>

TRUESWELL, J. C.; TANENHAUS, M. K. Toward a lexicalist framework for constraint-based syntactic ambiguity resolution. In: CLIFTON; FRAZIER; RAYNER (Ed.). *Perspectives on Sentence Processing*. Hillsdale, NJ: LEA Press, 1994. p. 155-179.

VOS, S. H.; FRIEDERICI, A. D. Intersentential syntactic context effects on comprehension: The role of working memory. *Cognitive Brain Research*, Bethesda, v. 16, n. 1, p. 111-122, 2003. [https://doi.org/10.1016/S0926-6410\(02\)00226-4](https://doi.org/10.1016/S0926-6410(02)00226-4)

WANNER, E.; MARATSOS, M. An ATN approach to comprehension. In: HALLE, M. BRESNAN, J.; MILLER, G. A. (Ed.). *Linguistic Theory and Psychological Reality*. Massachusetts: MIT Press, 1978.

WARREN, T.; GIBSON, E. Effects of NP type in reading cleft sentences in English. *Language and Cognitive Processes*, Taylor & Francis Online, v. 20, p. 751-767, 2005.

YANG, F.; MO, L.; LOUWERSE, M. M. Effects of local and global context on processing sentences with subject and object relative clauses. *Journal of Psycholinguistic Research*, Springer Link, v 42, n. 3, p. 227-237, 2013.

APPENDIX A

(1)

[SRC] O banqueiro que irritou o advogado joga tênis todo sábado.

[ORC] O banqueiro que o advogado irritou joga tênis todo sábado.

(2)

[SRC] A criança que achou a babá ria alegremente da brincadeira.

[ORC] A criança que a babá achou ria alegremente da brincadeira.

(3)

[SRC] O piloto que elogiou a aeromoça a chamou para um encontro.

[ORC] O piloto que a aeromoça elogiou a chamou para um encontro.

(4)

[SRC] O atacante que odiava o zagueiro agrediu o rival.

[ORC] O atacante que o zagueiro odiava agrediu o rival.

(5)

[SRC] O executivo que amava a secretária chamou até o porteiro para festa.

[ORC] O executivo que a secretária amava chamou até o porteiro para festa.

(6)

[SRC] O médico que ignorava a enfermeira dirigia um conversível vermelho.

[ORC] O médico que a enfermeira ignorava dirigia um conversível vermelho.

(7)

[SRC] O mecânico que namorava a garçonete sempre a convidava para almoçar.

[ORC] O mecânico que a garçonete namorava sempre a convidava para almoçar.

(8)

[SRC] A escritora que elogiou o fotógrafo trabalhava para uma revista de arte.

[ORC] A escritora que o fotógrafo elogiou trabalhava para uma revista de arte.

(9)

[SRC] O ladrão que assustou o policial já tinha assaltado três casas naquela semana.

[ORC] O ladrão que o policial assustou já tinha assaltado três casas naquela semana.

(10)

[SRC] a editora que aborreceu a escritora demitiu toda a equipe.

[ORC] a editora que a escritora aborreceu demitiu toda a equipe.

(11)

[SRC] O prisioneiro que atacou o carcereiro provocou a rebelião.

[ORC] O prisioneiro que o carcereiro atacou provocou a rebelião.

(12)

[SRC] O coreógrafo que admirava a dançarina deu a ela o papel principal.

[ORC] O coreógrafo que a dançarina admirava deu a ela o papel principal.

(13)

[SRC] A lavradora que desprezava o fazendeiro ligou para os jornais e reclamou.

[ORC] A lavradora que o fazendeiro desprezava ligou para os jornais e reclamou.

(14)

[SRC] A professora que criticou o estudante ficou um pouco envergonhada.

[ORC] A professora que o estudante criticou ficou um pouco envergonhada.

(15)

[SRC] A cliente que desagradou o psicólogo não voltou mais às sessões.

[ORC] A cliente que o psicólogo desagradou não voltou mais às sessões.

(16)

[SRC] O jogador que gostava do carregador foi chamado para um torneio profissional.

[ORC] O jogador que o carregador gostava foi chamado para um torneio profissional.

(17)

[SRC] O caçador que viu o guarda saiu correndo pela floresta.

[ORC] O caçador que o guarda viu saiu correndo pela floresta.

(18)

[SRC] O historiador que criticou o estudante se sentiu mal por aquela situação.

[ORC] O historiador que o estudante criticou se sentiu mal por aquela situação.

(19)

[SRC] A atriz que convocou o diretor pediu o papel principal do novo filme.

[ORC] A atriz que o diretor convocou pediu o papel principal no novo filme.

(20)

[SRC] O lobo que atacou a serpente feriu gravemente o animal.

[ORC] O lobo que a serpente atacou feriu gravemente o animal.

A aceitabilidade da anáfora logofórica em português brasileiro

The acceptability of logophoric anaphor in Brazilian Portuguese

Flávia Gonçalves Calaça de Souza

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil

flavia.ufpb@hotmail.com

Rosana Costa de Oliveira

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil

rosana.ufpb@gmail.com

Judithe Genuíno Henrique

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil

judithegh@gmail.com

Resumo: Este trabalho buscou investigar se as estruturas logofóricas são aceitáveis ou não para os falantes do português brasileiro (PB). Diversas abordagens sintáticas e semânticas têm sido usadas para explicar a logoforicidade, e uma forte discussão teórica é sobre qual tipo de abordagem melhor dá conta do fenômeno. Autores como Reinhart e Reuland (1993) apontam que a anáfora logofórica existe separadamente das anáforas sintáticas na Gramática Universal. As anáforas sintáticas são guiadas por fatores sintáticos como localidade e c-comando. Já a logófora pode ou não observar essas condições sintáticas. Esses estudos também demonstram que, nos casos em que há logoforicidade, a anáfora e o pronome são intercambiáveis, ou seja, tanto um quanto o outro podem ser usados na mesma posição na frase. Com base nesses pressupostos teóricos, examinamos a aceitabilidade de construções logofóricas em PB com a finalidade de saber se elas estão presentes na gramática dos falantes dessa língua. Foi realizado um experimento *off-line* de julgamento de aceitabilidade, visto que, segundo Gibson e Fedorenko (2013), a intuição

do linguista não é suficiente para averiguar tal aceitabilidade. O presente experimento mostrou que a anáfora logofórica é licenciada em PB, o que aponta que os sujeitos veem a anáfora logofórica e o pronome logofórico como diferentes. Não houve resultado significativo para o pronome, o que não nos permite apontar que não são licenciados em PB.

Palavras-chave: julgamento de aceitabilidade; logoforicidade; anáfora; pronome.

Abstract: This study sought to investigate whether these structures are acceptable or not for the speakers of Brazilian Portuguese. Both syntactic and semantic approaches have been used to explain the logophoricity and a strong theoretical discussion is about what kind of best approach to account for the phenomenon. Much of the literature on the subject Reinhart and Reuland (1993) points out that logophoric anaphora exists separately from syntactic anaphora in Universal Grammar. The syntactic anaphora is guided by syntactic factors such as location and c-command. Already logophor may or may not notice these syntactic conditions. These studies also demonstrate that where there logophoricity, the pronoun and anaphora are interchangeable, so, either one or the other can be used in the same position in the sentence. Based on these theoretical assumptions, we examine the acceptability of logophorics buildings in PB in order to know if they are present in the grammar of the speakers of that language. An offline experiment judgment of acceptability since that according to Gibson and Fedorenko (2013) the linguistics intuition is not enough to verify such acceptability. The following experiment showed the results were significant for logophoric anaphora, which shows that a degree I anaphora and logophoric pronoun as different. There was no significant result for the pronoun, which does not allow us to point out that are not licensed in PB.

Keywords: judgment of acceptability; logophoricity; anaphora; pronoun.

Recebido em: 10 de dezembro de 2016

Aprovado em: 4 de abril de 2017

1 Introdução

A Teoria da Ligação (*Binding Theory*) (CHOMSKY, 1981) busca estudar as relações de correferência que existem entre as expressões nominais e pronominais, dentro do escopo da sentença. Para explicar tais relações, Chomsky (1981) prediz três princípios que regem essas relações, a saber:

(1) Princípio A – Uma anáfora deve ser presa no seu domínio de ligação.¹

Princípio B – Um pronome deve ser livre no seu domínio de ligação.

Princípio C – Uma expressão R – referencial deve ser livre.²

O Princípio A, que se refere às anáforas reflexivas e recíprocas (*se, um ao outro*), prediz que uma anáfora deve ser presa em seu domínio de ligação, isto é, estar ligada por um antecedente que esteja c-comando e em seu domínio local. Compreende-se por c-comando a noção sintática de liame entre dois constituintes. Observemos o exemplo abaixo:

(2) [O neto do João_k]³ se_{i/*k} adora.

Neste exemplo, o DP “O neto do João” é o antecedente da anáfora “se”, assim o DP “O neto do João” está c-comando a anáfora “se”.

O Princípio B refere-se aos pronomes, este afirma que todo pronome deve ser livre no seu domínio de ligação, isto é, não pode haver a presença do antecedente e do pronome na mesma sentença. Vejamos os exemplos abaixo:

(3) a. [Penha_i acha que [Bruna_j vestiu ela_i com o vestido.]]
b. *[Penha_i acha que [Bruna_j vestiu ela_i com o vestido.]]

¹ Entendemos por domínio de ligação o limite que apresenta a anáfora e seu antecedente.

² Binding Principles (CHOMSKY, 1981)

A. An anaphor must be A-bound in its binding domain.

B. A pronominal must be A-free in its binding domain.

C. An R-expression must be A-free.

³ Índice de indexação.

A sentença (3a) é absolutamente plausível, pois o antecedente está fora do domínio do pronome, ao contrário da sentença (3b), que é agramatical, pelo fato de termos o pronome e seu antecedente no mesmo domínio, conforme defende a Teoria da Ligação.

E por fim, porém não menos importante, temos o Princípio C. Este faz menção às expressões referenciais, que devem ser livres em qualquer contexto de ocorrência, pois possuem autonomia referencial e, assim, não precisam ser ligadas. Observemos o exemplo abaixo:

- (4) O João observou o Pedro na sala.

Em alguns casos, a anáfora não está restrita às condições de ligação propostas por Chomsky (1981), como mostra Reinhart e Reuland (1993). Estes autores denominam esse tipo de anáfora de logofórica. Vejamos:

- (5) Maria contava cinco pessoas na cozinha, além de si mesma/dela.
(6) Carlos viu uma blusa perto de si/dele.

Nos exemplos acima, a anáfora está em um PP adjunto, não sendo um argumento exigido pelo verbo. Segundo Reinhart e Reuland (1993), Chomsky (1981, 1986) não explica os contextos em (5) e (6), justamente pelo fato da anáfora e seu antecedente não serem coargumentos. Essas anáforas não estão de acordo com o previsto pela Teoria da Ligação, pois esta não explica a correferência com antecedentes fora do domínio de ligação.

A partir destes e de outros contraexemplos, Reinhart e Reuland (1993) observam as anáforas a partir da noção de argumentos, apontando que a anáfora é logofórica quando não está em posição argumental. A anáfora sintática, como o próprio nome aponta, é guiada por fatores sintáticos como localidade e c-comando, e a anáfora logofórica relaciona-se com os fatores discursivos. Como nos exemplos que seguem:

- (7) João se vangloriava de que o prefeito convidou Ana e ele mesmo para um café da manhã.
(8) Vânia encontrou tempo para verificar que, além dela mesma, havia um outro funcionário que tinha sido demitido.

Nas sentenças acima nota-se que a previsão de Chomsky (1981; 1986) para as anáforas não podem ser verificadas, uma vez que em ambas as sentenças a anáfora não está no domínio de ligação. Zribi-Hertz (1989) fornece cerca de 130 exemplos, como os apresentados acima, de sentenças em que a anáfora não se encontra em posição argumental.

Reinhart e Reuland (1993) ainda afirmam que é possível haver um pronome logofórico sempre que houver uma anáfora logofórica, como em (9) e (10):

- (9) Carlos estava chateado quando Júlia colocou Bruno e ele mesmo em perigo no paraquedismo.
- (10) Carlos estava chateado quando Júlia colocou Bruno e ele em perigo no paraquedismo.

O tipo de estrutura das sentenças (9) e (10) é bastante discutido por Pollard e Sag (1992) e Reinhart e Reuland (1993), visto que é consistente com o fato de que em inglês o pronome é uma forma aceitável para ser alternada com a anáfora, mantendo a mesma interpretação.

O presente trabalho tem como foco verificar a aceitabilidade da anáfora logofórica e do pronome logofórico em PB, isto é, nosso objetivo foi averiguar se os sujeitos julgam as sentenças com anáfora logofórica e pronome logofórico como aceitáveis ou inaceitáveis de acordo com sua intuição.

Gibson e Fedorenko (2013) apontam que, apesar das intuições do pesquisador serem úteis para diversos estudos, em muitos outros não se pode ter uma confiabilidade alta apenas através da intuição, exigindo um controle metodológico mais rigoroso a fim de detectar sutilezas e distinguir entre diferentes posições teóricas. O autor ainda afirma que um grande problema com o campo de estudo da sintaxe é que muitos artigos não incluem a evidência experimental como apoio de suas hipóteses de pesquisa. Também é recomendado pelo autor que os estudos reúnam evidências quantitativas para que haja uma melhoria para este campo de estudo. Levando em conta tal importância, nesta pesquisa demos um enfoque ao método quantitativo de Julgamento de Aceitabilidade, em que os sujeitos fizeram o julgamento das sentenças com anáfora logofórica e pronome logofórico com o intuito de apresentar dados confiáveis sobre a aceitabilidade dessas sentenças.

O experimento *offline* de julgamento de aceitabilidade realizado, teve como objetivo central observar a aceitabilidade de anáforas consideradas logofóricas, especificamente da anáfora “ele mesmo” em português brasileiro (PB), com a finalidade de compará-la ao pronome “ele”, para compreender se, assim como no inglês, esse fenômeno está presente na gramática dos falantes do PB. Por hipótese, com base na Teoria da Reflexividade de Reinhart e Reuland (1993), essas estruturas seriam aceitáveis e não violariam as condições de ligação por não estarem restritas a elas.

É importante destacar que existe uma escassez de estudos em português brasileiro sobre a resolução de logofóricas e os processos que ela exige. Assim, nosso trabalho, além de verificar a aceitabilidade das estruturas logofóricas em PB, buscou fazer um levantamento bibliográfico de estudos realizados em outras línguas sobre esse tema que apontam que a anáfora logofórica existe separadamente da anáfora sintática na Gramática Universal (GU).

2 A logoforicidade em Linguística

Apesar de em alguns casos as relações entre anáforas, pronomes e os seus antecedentes serem estabelecidas obedecendo a restrições da Teoria da Ligação (CHOMSKY, 1981), alguns estudos (REINHART; REULAND, 1993; ZRIBI-HERTZ, 1989) mostram que essas formas podem ser interpretadas de outras maneiras. Como apontado acima, a logoforicidade ocorre nesses casos e tem sido caracterizada e definida a partir de diversas perspectivas. Uma dessas perspectivas é não estar sujeita às restrições de c-comando e localidade impostas pela Teoria da Ligação ao possuir um *status* logofórico.

A seguir, serão delineadas, em mais detalhes, as abordagens de Chomsky (1981), Reinhart e Reuland (1993) e Zribi-Hertz (1989), com a finalidade de fazer um levantamento descritivo desse tema tão pouco pesquisado. Algumas dessas abordagens observam a anáfora logofórica a partir da configuração de c-comando, ou seja, tanto limitando e definindo por meio das restrições de natureza sintática, quanto observando o discurso e a semântica antes da sintaxe. Na abordagem de Zribi-Hertz (1989), por exemplo, a prioridade teórica é dada aos fatores discursivos. Caso estes não deem conta, o domínio sintático é procurado.

2.1 Chomsky (1981)

A Teoria da Ligação formulada por Chomsky (1981) descreve as restrições sintáticas sobre as anáforas, pronomes e expressões referenciais. Ele postulou três princípios de vinculação: A, B e C. O princípio A prediz que uma anáfora deve ser ligada ao antecedente que está em seu domínio, ou seja, deve estar ligada ao antecedente que está na oração mínima que contém a anáfora. Além da presença de um antecedente local, o princípio A também prediz que o antecedente c-comanda a anáfora. Nas sentenças abaixo, temos dois exemplos de anáforas vinculadas a seu antecedente local:

(11) [A prima de Joana]_i se_i acha muito bonita.

(12) O Carlos_i machucou ele mesmo_i enquanto cozinhava.

Já em (13) e (14), a seguir, vemos o exemplo de um DP que não está c-comando as anáforas *se* e *ele mesmo*, pois, de acordo com a Teoria da Ligação, a anáfora *se* não pode se referir a *Joana* e a anáfora *ele mesmo* não pode se referir a *Carlos*, tornando as sentenças agramaticais.

(13) *A prima de [Joana]_i se_i acha muito bonita.

(14) *O Carlos_i machucou Vivian e ele mesmo_i enquanto cozinhava.

O princípio B prevê que o pronome esteja livre em seu domínio de vinculação. Na sentença (15) abaixo, o pronome *ela* não pode ter como antecedente o DP *Maria*, pois está em seu domínio de vinculação.

(15) *A Maria_i adora ela_i.

O princípio B, portanto, impõe que o pronome não pode ser ligado ao seu antecedente local, distintamente do princípio A, que afirma que a anáfora tem que estar ligada a esse antecedente.

Já o princípio C determina que as expressões referenciais devem estar livres em qualquer contexto sintático, isto é, não podem estar ligadas. Em (16), a expressão-R *o animal* não possui antecedente na sentença, o que a torna gramatical. As expressões-R não necessitam de um elemento antecedente para extrair seu significado, ou seja, são independentes referencialmente, o que as diferencia das anáforas e dos pronomes, que possuem dependência referencial.

(16) O animal fez um grande barulho.

Ainda de acordo com a Teoria da Ligação, as anáforas e os pronomes estão em distribuição complementar. Isto é, um elemento pode estar na sentença somente onde o outro não pode. Assim, para que a sentença seja considerada gramatical, esses elementos devem estar em ambientes opostos. Uma anáfora deve ser ligada e um pronome deve ser livre. Portanto, se uma anáfora possuir um antecedente dentro da mesma oração que a inclui, a sentença é gramatical. Mas se nessa mesma sentença houver um pronome em vez de uma anáfora, seria agramatical. No entanto, vários autores têm demonstrado que esta distribuição complementar nem sempre é possível de se constatar, pois sentenças como (17) são consideradas gramaticais.

(17) Mateus_i sabia que ninguém gosta de João e dele mesmo_i/dele_i por causa de suas personalidades.

A proposta de Reinhart e Reuland (1993), que veremos a seguir, aponta que as anáforas como em (17) sejam observadas a partir da noção de argumentos e, caso não sejam coargumentos do mesmo predicado, permanecerão sendo consideradas gramaticais. A ideia fundamental é que a distribuição complementar não exista em sentenças como em (17), que contém uma anáfora logofórica.

2.2 Reinhart e Reuland (1993)

A Teoria da Reflexividade proposta por Reinhart e Reuland (1993) surge a partir da necessidade de uma reformulação na Teoria da Ligação (CHOMSKY, 1981) justificada pelos muitos problemas empíricos que apresentaram os princípios A e B dessa teoria. Os autores também questionam a distribuição complementar entre anáforas e pronomes, levantando vários exemplos em que essa complementariedade falha. Essa teoria ainda deixa clara a diferença quando há ligação e quando há correferência entre as formas anafóricas.

Para tanto, os autores propõem retornar à interpretação dos predicados reflexivos nas línguas naturais em que o verbo tem uma ligação estreita com o fenômeno da reflexivização. A partir disso, a reflexivização seria então analisada como uma propriedade dos predicados e não como uma propriedade da anáfora, ou seja, a distribuição das formas anafóricas

é feita a partir das propriedades semânticas das formas anafóricas, nesse caso, a propriedade de reflexivizar um predicado (*inherent properties*). É assumida, então, a seguinte definição de anáfora:

Anaphors (of both the SE and the SELF type) are referentially defective NPs, which entails, for example, that they cannot be used as demonstratives, referring to some entity in the world (though it does not entail that they must be bound variables). (REINHART; REULAND, 1993, p. 658).⁴

A partir dessa definição e das funções de ambos os tipos de anáfora, é proposta uma tipologia das expressões anafóricas. As anáforas SE e as anáforas SELF formam o mesmo grupo em relação à propriedade de dependência referencial, e as anáforas SE e os pronomes agrupam-se por não terem a função de reflexivizar. O resumo dessa tipologia segue na tabela abaixo:

TABELA 1 – Proposta apresentada por Reinhart e Reuland (1993) para a distribuição das anáforas e pronomes

	SELF	SE	PRONOUN
Reflexivizing function	+	-	-
R(eferential Independence)	-	-	+

Fonte: Reinhart e Reuland (1993)

De acordo com a tabela acima, as anáforas SELF têm a propriedade de reflexivizar um predicado (+ função reflexiva), mas não têm independência referencial. Os autores consideram que essa função reflexiva SELF é sensível às condições de ligação propostas por eles. Já as anáforas SE não têm independência referencial como também não

⁴ “Anáforas (de ambos os tipos SE e SELF) são NPs referencialmente defeituosos, o que significa, por exemplo, que eles não podem ser utilizados como demonstrativos, referenciando alguma entidade no mundo (embora isto não signifique que eles devem ser variáveis ligadas)”.

são responsáveis por tornar um predicado reflexivo (- função reflexiva), por serem apenas marcas para predicados inerentemente reflexivos. Os pronomes, por sua vez, não reflexivizam predicados (- função reflexiva), mas, ao contrário das anáforas, têm independência referencial por possuírem traços de pessoa, gênero e número – não obrigatoriamente todos.

É importante destacar que, para os autores, os dois tipos de anáforas citadas podem ser usadas logoforicamente. Entretanto, devido à escassez de estudos que façam a distinção das funções gramaticais e logofóricas da anáfora, há muitos equívocos com relação à distribuição sintática das anáforas e os termos local e longa distância, já que, se usadas logoforicamente, os dois tipos de anáforas podem ocorrer em todos os tipos de distância. Deste modo, os reflexivos logofóricos não possuem uma regra específica, mas devem ser considerados gramaticais. No entanto, os autores apontam que uma propriedade da anáfora logofórica é não ser c-comandada (REINHART; REULAND, 1993, p. 660).

Reinhart e Reuland (1993) discutem que é possível haver um pronome sempre que houver uma anáfora logofórica. Ao afirmar isso, os autores não apresentam uma justificativa das razões de preferência entre a anáfora logofórica e o pronome. Uma abordagem sugerida, em nota, pelos próprios autores, é a de Ariel (1990), que leva em conta os candidatos disponíveis mais acessíveis no discurso.

A respeito da função gramatical, afirmam os autores que os domínios são apenas dois: local e longa distância. O domínio local corresponde ao domínio de reflexividade no qual a anáfora SELF obrigatoriamente reflexiviza o predicado. Esse domínio é regulado pelas Condições A e B reformuladas por eles e as quais descreveremos adiante. Já o domínio de longa distância refere-se ao domínio de ligação das anáforas SE que seguem o padrão dos pronomes e, por isso, são regidos pela Condição B.

As Condições A e B devem ser lidas como condicionais e são assim definidas:

- (i) Condição A: um predicado sintático marcado reflexivamente é reflexivo.
- (ii) Condição B: um predicado semântico reflexivo é reflexivamente marcado.

A Condição A aplica-se a predicados que são marcados reflexivamente por vias sintáticas e a Condição B inerentemente (semanticamente) a reflexivos. Entende-se por predicado sintaticamente reflexivo como sendo aquele que tem dois dos seus argumentos coindexados (com os traços de gênero, número e pessoa indistintos) e por predicado reflexivamente marcado aquele ou que é lexicalmente reflexivo (o que é marcado pela presença de uma anáfora SE) ou cujo um dos seus argumentos indexados é uma anáfora SELF. Logo, não é a anáfora em si, mas a marcação reflexiva que confere que seja interpretado reflexivamente.

Sendo assim, para Reinhart e Reuland, as Condições A e B estão relacionadas não com a configuração sintática de c-comando, isto é, com a possibilidade de ligação com o antecedente disponível no mesmo domínio da anáfora ou pronome, mas com a propriedade intrínseca do predicado de ser ou não inerentemente reflexivo.

Com a finalidade de observarmos como as Condições A e B se aplicam, vejamos os exemplos abaixo:

- (18) a. **John_i likes him_i.*
 b. *John_i likes himself_i.*
 c. *John_i said Ann likes him_i.*⁵

Observe que em (18a) *John* e *him* são coargumentos do predicado *likes*. Esse predicado sintático não será reflexivo porque ele não foi marcado pela anáfora do tipo SELF, que corresponde à anáfora que reflexiviza predicados, mas foi substituído por um pronome, tornando a frase agramatical. Um predicado como *like* estaria submetido à condição A – um predicado que seria sintaticamente marcado como reflexivo porque um dos seus argumentos seria uma anáfora SELF. Em (18b), a frase é gramatical porque a anáfora SELF está na sentença como marcador reflexivo e está condizente com a Condição A. Já em (18c), *John* é coargumento de *said* e *him* é coargumento de *likes*. Portanto, o predicado não é reflexivo, já que não há coindexação entre os argumentos de um mesmo predicado e, por isso, não é guiado pela Condição B. Em

⁵ *John gosta dele.

John gosta dele mesmo.

John disse que Ann gosta dele.

(18c), os predicados não são semanticamente reflexivos e, por isso, não são reflexivamente marcados.

Um dos problemas apontados pelos autores sobre a Condição A da Teoria da Ligação (CHOMSKY, 1981) é a ocorrência de anáforas livres de seu domínio, como mostrado por Kuno (1987) e Zribi-Hertz (1989) com anáforas SELF de primeira, segunda e terceira pessoas. Vejamos:

(19) *Max boasted that the queen invited Lucie and himself for a drink.*

(20) **Max boasted that the queen invited himself for a drink.*⁶

Reinhart e Reuland (1993) citam que, para Zribi-Hertz (1989), as anáforas de terceira pessoa ligadas à longa distância violam o princípio A e têm o uso conhecido como logofórico. Entretanto os autores argumentam que não é possível concluir que os contextos em que há ponto de vista⁷ permitem uma violação da Condição A. Para Reinhart e Reuland (1993), não se trata apenas de uma questão discursiva, mas, sobretudo, de uma questão estrutural, visto que em (20) a anáfora SELF é um argumento de seu predicado e, por isso, o predicado é marcado reflexivamente. Em (19), não há predicado marcado reflexivamente, pois o argumento de *invited* não é apenas *himself* (anáfora SELF), mas *queen*, com o qual *himself* não compartilha traços de gênero e, por isso, não pode ser a ele coindexado.

A visão desenvolvida pelos autores, a partir da distinção entre função gramatical e logofórica, mostra que a Condição A apenas é aplicada à anáfora em posição argumental. Ou seja, as anáforas SELF que ocorrem fora dessa posição (logofóricas) estão isentas dessa condição.

Ainda nesta discussão, os autores afirmam que a distinção entre uso anafórico e uso logofórico do SELF é supérfluo, porque sintaticamente existe apenas um tipo de anáfora, cujo uso é regido pela Condição A. Isso exclui, de imediato, a anáfora SELF, em que não há predicados reflexivos como aplicáveis a essa Condição. Portanto, a logoforicidade não é codificada na sintaxe e o logóforo não tem que

⁶ Max se gabou que a rainha convidou Lucie e ele mesmo para uma bebida.

Max se gabou que a rainha convidou ele mesmo para uma bebida.

⁷ Tem sido caracterizada na literatura a noção de ponto de vista para indicar que a correferência da anáfora/pronome com o autofalante, nos termos de Kuno (1987) cujo discurso, pensamentos, sentimentos estão sendo relatados.

ser ligado/coindexado a um antecedente. A sua relação pode ser de correferência (REINHART; REULAND, 1993, p. 673). Essa postura entra em discordância com Sells (1987), que afirma que os logóforos, especialmente os de perspectiva, são ligados.

Há apenas um caso em que a Condição A permite o uso logofórico da anáfora SELF: quando esta não ocupa uma posição argumental e, conseqüentemente, não marca o predicado como reflexivo.

Como afirmado acima, a Condição A governa a anáfora. Além disso, também é postulado pela Teoria da Reflexividade que o pronome não é excluído em certos contextos nos quais há a predição de uma anáfora ocorrer. Um exemplo desse contexto é quando há um *NP picture*, como em (21), no qual se prevê um mecanismo adicional para justificar a ocorrência do pronome.

(21) *Mary saw a Picture of herself*⁸

Além disso, é importante enfatizar que uma propriedade da anáfora logofórica citada pelos autores é que ela não tem que ser c-comandada pelo seu antecedente. Dessa forma, os exemplos de *NP picture* são facilmente aceitos, pois há correferência e não há ligação.

2.3 Zribi-Hertz (1989)

O artigo de Zribi-Hertz publicado em 1989 faz um levantamento detalhado de ocorrências de pronomes reflexivos do inglês que são compatíveis com a Teoria da Ligação chomskyana e observa que algumas dessas anáforas violam o princípio A dessa teoria, o que mostra, segundo ela, que se baseia em um corpo relativamente pequeno de dados e não é suficiente para provar que a distribuição dos pronomes do uso cotidiano é tão restrita quanto a previsão feita pelo princípio A.

Esses pronomes que violam tal princípio, segundo a autora, traçam uma linha clara entre sintaxe e discurso e mostram que o princípio A de Chomsky é completo apenas se for uma teoria interna da frase, sem levar em consideração o discurso, mas incompleta porque ignora um componente integral da gramática de reflexivos e assim não dá conta de vários conjuntos de dados.

⁸ Mary viu uma foto dela mesma.

A discussão central do estudo de Zribi-Hertz é compreender onde a linha deve ser traçada, com relação à anáfora, entre sintaxe e discurso. É feito um levantamento sobre as propriedades sintáticas e semânticas dos pronomes reflexivos em inglês a fim de confrontá-las com um *corpus* de textos do inglês. Após a análise do *corpus*, a autora esclarece que os pronomes reflexivos do inglês podem ser ligados a longa distância e podem violar várias restrições estruturais. A partir disso, é proposta uma relação entre a gramática dos reflexivos localmente ligados e a gramática discursiva dos reflexivos ligados a longa distância.

Entre os dados problemáticos que surgiram após a Teoria da Ligação, alguns põem em xeque a complementariedade entre anáforas e pronomes. Alguns deles são: NPs *Picture*, posições de genitivo, alguns *Prepositional phrase* (PPs) e contextos enfáticos, que são discutidos por Warshawsky (1965), Ross (1970), Cantrall (1974), Kuno (1987) e outros autores. Vejamos:

- (22) a. *They thought that [pictures of {them / themselves}] would be on sale.*
 b. *We thought that [John's pictures of {us / ourselves}] would be on sale.*
- (23) *John said that there was a picture of {him / himself} in the post office.*
- (24) a. *(Mary thought that) [a picture of {you / yourself}] would be nice on the wall.*
 b. *(Mary thought that) [a picture of {me / myself}] would be nice on the wall.*
- (25) *They heard the stories about {them / themselves}.*⁹

⁹ (22) a. Eles achavam que [fotos {deles / deles mesmos}] seriam colocadas à venda.
 b. Nós pensamos que [as fotos de John de {nós / nós mesmos}] seriam colocadas à venda.

(23) João disse que havia uma imagem {dele / dele mesmo} na estação de correios.

(24) a. (Mary pensou que) [uma imagem {sua/ de si mesmo}] seria bom na parede.

b. (Mary pensou que) [uma imagem {sua / de si mesma}] seria bom na parede.

(25) Eles ouviram as histórias sobre eles {Eles/eles mesmos}.

A partir desses e de outros contraexemplos, alguns autores apontam ora para uma mudança ou alteração na tipologia a fim de abarcá-los dentro da teoria estrutural, ora é sugerido que essas ocorrências estejam fora do âmbito da sintaxe. Contudo, é argumentado neste artigo que uma teoria gramatical dos pronomes reflexivos do inglês não pode ser completa sem um componente discursivo.

A autora cita a discussão de Cantrall (1969) sobre a hipótese de que, em inglês, onde é possível alternar pronomes e anáforas no mesmo contexto estrutural, a escolha de um ou do outro expressa uma escolha de ponto de vista narrativo, como também a de Kuroda (1973), que sugere que a opção reflexiva de *zibun* está correlacionada com o estilo narrativo não reportivo por meio do qual o autor relata os eventos “de dentro” (algum personagem) ao contrário do seu próprio ponto de vista.¹⁰

Essa discussão de Cantrall (1969) é possível de ser observada no exemplo abaixo:

- (26) a. *The women_i were standing in the background, with the children behind them_i.*
b. *The women_i were standing in the background, with the children behind themselves_i.*¹¹

As sentenças 26a e 26b não possuem o mesmo conteúdo informativo, pois em 26b as crianças estão localizadas “atrás das mulheres”, do ponto de vista interno do protagonista discursivo (as mulheres), e em 26a, as crianças estão “atrás das mulheres”, do ponto de vista do falante. O contraste está relacionado à opção do ponto de vista, que pode ser de um protagonista do discurso ou do falante e, conseqüentemente, como aponta Cantrall (1969), nas escolhas de uma terceira pessoa das anáforas pronominais e pronomes, não há um condicionamento estrutural.

A proposta da Teoria da Ligação de Chomsky não faz correlação com as propriedades semânticas acima e é assumido que elas não pertencem à sintaxe e são derivadas de suas propriedades estruturais. No

¹⁰ O ponto de vista também é o conceito central das análises dos pronomes em Kuno (1972, 1983, 1987).

¹¹ a. As mulheres estavam em pé ao fundo, com as crianças atrás delas.

b. As mulheres estavam em pé ao fundo, com as crianças atrás delas mesmas.

entanto, a autora, apesar de concordar que o ponto de vista não pertence à sintaxe, discorda de sua derivação a partir das propriedades estruturais: “*It is the structural properties of pronouns that are, in a sense, derived from their discourse properties*”.¹² (ZRIBI-HERTZ, 1989, p. 705)

Para justificar sua afirmação, são dados os exemplos em (27), em que há um contraste semântico:

- (27) a. *John_i hid the book behind himself_i.*
 b. *John_i hid the book behind him_i.*¹³

Para Zribi-Hertz, a relação espacial entre John e o livro é entendida como mais direta em 27a do que em 27b porque, nesta última, John escondeu o livro em algum lugar atrás da linha dos ombros, enquanto que em 27a John escondeu o livro muito perto dele, provavelmente em contato com seu corpo.

É argumentado, então, que o “sujeito de consciência” não é uma variação semântica do sujeito sintático, mas que se destaca como um conceito linguístico pertencente à gramática do discurso e que este, em vez do sujeito sintático, é o conceito relevante para a gramática dos reflexivos ligados à longa distância no inglês.

O conceito “sujeito de consciência” é uma categoria da gramática do discurso, sendo similar ao conceito de logoforicidade proposto por Kuno (1987) e Clements (1975). Como afirma a autora, o “sujeito de consciência” é uma propriedade semântica atribuída a um referente cujos pensamentos ou sentimentos, opcionalmente expressos no discurso, são transportados por uma parte do discurso. Esse “sujeito de consciência”, ainda nas palavras da autora, é geralmente entendido como [+ humano].

A autora diz que os pronomes reflexivos pertencem à gramática do discurso, da qual a gramática da sentença é um subdomínio específico. Por meio dessa conclusão, a autora sugere que as restrições estruturais podem ter a sua motivação discursiva.

Uma afirmação importante colocada pela autora é a de que as relações anafóricas apontadas em seu estudo pertencem à gramática do

¹² “São as propriedades estruturais dos pronomes que são, em certo sentido, derivadas de suas propriedades discursivas.”

¹³ a. John escondeu o livro atrás dele mesmo.

b. John escondeu o livro atrás dele.

discurso, visto que as restrições estruturais entram em jogo quando o princípio do discurso deixa de ser relevante.

Portanto, a proposta de Zribi-Hertz afirma que qualquer teoria estrutural da anáfora deve ser preenchida pelos princípios de discurso, pois, para dar conta desse fenômeno, é preciso considerar as noções de estrutura do discurso.

Essa proposta de a gramática da sentença (sintaxe) ser uma subparte da gramática do discurso autoriza os pronomes reflexivos a violarem o Princípio A de Chomsky, caso certas propriedades discursivas sejam satisfeitas. Isso torna essa proposta diferente da de Reinhart e Reuland (1993), pois, para esses últimos, a sintaxe opera prioritariamente e, caso não seja suficiente, a semântica e o discurso estão livres para atuar.

3 A logoforicidade em Psicolinguística

Apesar das diversas abordagens sobre o fenômeno, há poucos estudos em Psicolinguística Experimental que têm como foco principal compreender o processamento desse tipo de estrutura. Neste tópico mostraremos algumas abordagens no processamento intrassentencial e alguns experimentos desenvolvidos em Psicolinguística.

No âmbito da Psicolinguística Experimental intrassentencial, há uma discussão sobre o processamento *on-line* de sentenças que possuem uma única oração ser restrito exclusivamente à atuação dos Princípios de Ligação postulados por Chomsky (1981) nas primeiras etapas do processamento. Nessa problemática existe um forte debate sobre o momento de atuação dos princípios de ligação de Chomsky, no curso do processamento, o que tem se desdobrado em predições diferenciadas com os modelos que postulam o quanto o processamento é afetado pelos seus antecedentes, a saber, os modelos: filtro inicial, filtro reversível e interativo.

Em linhas gerais, o primeiro modelo foi apresentado por Nicol e Swinney (1989), que afirma que a ligação entre o antecedente e a anáfora é restrita aos princípios de ligação de Chomsky (1981) tanto nos estágios iniciais do processamento quanto nos estágios que o seguem. O segundo modelo é defendido por Sturt (2003), Kennison (2003) e Leitão, Peixoto e Santos (2008) ao apresentar evidências de um filtro reversível em que esses princípios guiam o processamento nos estágios iniciais, mas podem ser violados posteriormente por vários fatores. Já Badecker e

Straub (2002) analisam seus dados interpretando que não só as restrições sintáticas atuam nesse primeiro estágio, mas também outras restrições de caráter discursivo. Esses últimos propõem o modelo interativo.

O modelo descrito por Nicol e Swinney (1989) mostra que as restrições propostas pela Teoria da Ligação são aplicadas nos estágios iniciais e subsequentes do processamento. De acordo com esse modelo, o antecedente para a anáfora é imediatamente escolhido de acordo com a teoria da ligação, já outros antecedentes são desconsiderados nos primeiros estágios do processamento e posteriormente.

Os autores analisaram o processamento de pronomes e anáforas que possuíam dois tipos de antecedentes: disponíveis e indisponíveis, segundo a Teoria da Ligação. Eles realizaram um experimento de *priming cross modal*¹⁴ (inter-modal) com a finalidade de observar a atuação das restrições de ligação. Nesse experimento os participantes ouviram as sentenças e fizeram uma decisão lexical após visualizarem uma palavra sonda que aparecia após as anáforas e pronomes.

Dessa forma, previa-se que os antecedentes disponíveis estruturalmente influenciariam o processamento da correferência nos estágios iniciais e que os antecedentes indisponíveis seriam imediatamente excluídos e desconsiderados na interpretação subsequente (Hipótese do Filtro Inicial).

Os resultados do estudo realizado por eles trouxeram evidências para a Hipótese do Filtro Inicial, pois houve efeito significativo para os antecedentes disponíveis estruturalmente e não houve efeito significativo para os antecedentes indisponíveis estruturalmente tanto para a anáfora quanto para o pronome. Isso indica que eles não foram levados em consideração no momento da resolução da correferência e há um filtro gramatical que determina quais candidatos podem se referir à anáfora e ao pronome.

Entretanto, Sturt (2003), Kennison (2003) e Leitão, Peixoto e Santos (2008) encontraram evidências contra o modelo do filtro de ligação inicial. Seus dados mostram que as restrições de ligação são aplicadas nos estágios iniciais, mas podem ser violadas posteriormente. O filtro seria, então, reversível, com a possibilidade de ser violado em um segundo momento do processamento por vários fatores.

¹⁴ Quando os estímulos apresentados no priming e no alvo são de modalidades diferentes.

O trabalho de Sturt (2003) mostrou que as restrições de ligação são relevantes no processamento da correferência no primeiro estágio do processamento. Os fatores discursivos influenciariam apenas tardiamente, o que justifica o efeito dos antecedentes indisponíveis encontrados no seu estudo.

Na mesma direção dos achados de Sturt (2003), Kennisson (2003) investigou a atuação do princípio B no processamento por meio da técnica de leitura automonitorada, capturando o processamento em dois estágios.

A autora fez um experimento utilizando os pronomes *her*, *his* e *him* com antecedentes disponíveis e indisponíveis, no qual os sujeitos leram mais lentamente o pronome na condição em que o antecedente indisponível estruturalmente era do mesmo gênero da retomada em comparação à condição em que o gênero era o mesmo.

A partir dos resultados encontrados, Kennisson (2003) sugere que os antecedentes disponíveis e indisponíveis estruturalmente são considerados durante a resolução correferencial, diferentemente da proposta de Nicol e Swinney (1989) de que apenas são considerados os antecedentes disponíveis estruturalmente.

Dialogam com os resultados de Kennisson (2003) e Sturt (2003) os achados de Leitão, Peixoto e Santos (2008) para o PB encontrados por meio de dois experimentos de leitura automonitorada. No experimento realizado pelos autores, foi investigado o processamento do pronome “ele” na posição de objeto. No primeiro experimento, não houve diferença significativa na leitura do segmento crítico (pronome), sugerindo-se que o princípio B bloqueou, no primeiro estágio, a possibilidade do “ele” se vincular ao sujeito.

Os tempos de leitura foram mais lentos quando o sujeito não disponível estruturalmente possuía os mesmos traços de gênero, número e animacidade do pronome do que nas condições em que esses traços não combinavam. Esses resultados sugerem que o processamento da correferência ocorre em dois estágios.

Ao segundo experimento foi acrescentado um preâmbulo com um antecedente disponível que combinava os mesmos traços com o pronome. A presença desse antecedente disponível tornou a leitura do pronome significativamente mais demorada do que no experimento 1, ou seja, não houve a influência dos antecedentes indisponíveis.

Por fim, Badecker e Straub (2002) propõem um modelo no qual várias restrições são atuantes no momento inicial do processamento.

Tanto sintáticas quanto discursivas são combinadas paralelamente, ou seja, as restrições de ligação atuam junto com outros tipos de informações (traços de gênero e número, foco discursivo). Neste entendimento, tanto os antecedentes disponíveis quanto os indisponíveis são relevantes no processamento correferencial desde o primeiro estágio do processamento.

No estudo de leitura automonitorada feito por eles, havia antecedentes tanto disponíveis quanto indisponíveis estruturalmente. Badecker e Straub (2002) encontraram resultados no segmento pós-crítico, em que foram lidos mais rapidamente quando o gênero da retomada combinava com o gênero do antecedente. Porém, a leitura foi mais lenta quando o gênero dos antecedentes indisponíveis e disponíveis combinavam com o gênero da retomada, o que mostra que os antecedentes indisponíveis também influenciaram no processamento das sentenças. Nesta linha, os autores propõem que os antecedentes indisponíveis interferem no processamento.

Entretanto, esses modelos citados acima não contemplam, especificamente, a discussão do processamento de anáforas logofóricas e pronomes logofóricos, apesar de citarem a possibilidade de haver certos tipos de estrutura em que há o uso da anáfora e não há a atuação dos princípios de ligação.

Em um outro estudo psicolinguístico, envolvendo especificamente a logoforicidade, Foraker (2003) usou a forma SELF logoforicamente para examinar se as informações discursivas utilizadas na interpretação desse logofórico são semelhantes às usadas na interpretação do pronome. Nesse estudo foi realizado um experimento de leitura automonitorada no qual foi controlada a distância entre o logofórico / pronome e seu antecedente, com três posições possíveis para o antecedente (*First-mentioned, Middle e Most Recent*). Nesse experimento, Foraker mostra que houve um efeito *Spill over* (após o logofórico / pronome), o qual revelou que logofóricos e pronomes foram processados de forma semelhante apenas nas condições em que o antecedente estava a uma distância longa ou intermediária em relação à retomada. Quando o antecedente foi mencionado mais recentemente, o logofórico foi lido mais rapidamente do que quando este estava em posição distante ou intermediária.

A sua proposta de explicação para os resultados encontrados nas condições com antecedente em distância longa e intermediária é a de que a resolução da correferência da anáfora logofórica parece usar o mesmo tipo de informação pragmática e discursiva usada para correferência

pronominal, já que, tanto para logofóricos quanto para pronomes foi encontrado o mesmo padrão de processamento. O conjunto de frases do experimento segue exemplificado em (28) e (29) abaixo:

(28) *Logophoric Reflexive*

First-mentioned: Megan wondered/ if Isaac had found out/ that Rick wanted to invite/ Sally and herself/ to the birthday party. Middle: Isaac wondered/ if Megan had found out/ that Rick wanted to invite/ Sally and herself/ to the birthday party. Most Recent: Rick wondered/ if Isaac had found out/ that Megan wanted to invite/ Sally and herself/ to the birthday party.

(29) *Pronoun*

*First-mentioned: Albert was upset/ when Debbie didn't care/ that Rachel had endangered/ Gordon and him/ on the climbing trip. Middle: Debbie was upset/ when Albert didn't care/ that Rachel had endangered/ Gordon and him/ on the climbing trip. Most Recent: Rachel was upset/ when Debbie didn't care/ that Albert had endangered/ Gordon and him/ on the climbing trip.*¹⁵

Ainda no inglês, Harris et al (2000) fez um estudo com EEG¹⁶ com a finalidade de verificar as características das respostas de ERP¹⁷ em frases com violação sintática e violação não sintática, já que na literatura

¹⁵ Reflexivos logofóricos.

Mencionado primeiro: Megan perguntou se Isaac tinha achado que Rick queria convidar Sally e ela mesma para a festa de aniversário. Médio: Isaac perguntou se Megan havia achado que Rick queria convidar Sally e ela mesma para a festa de aniversário. Mais recente: Rick perguntou se Isaac tinha achado que Megan queria convidar Sally e ela mesma para a festa de aniversário.

Pronome

Mencionado primeiro: Albert ficou chateado quando Debbie não se importou que Rachel tivesse ameaçado Gordon e ele na viagem de escalada.

Médio: Debbie ficou chateada quando Albert não se importou que Rachel tivesse ameaçado Gordon e ele na viagem de escalada.

Mais recente: Rachel ficou chateada quando Debbie não se importou que Albert tivesse ameaçado Gordon e ele na viagem de escalada.

¹⁶ Eletroencefalógrafo.

¹⁷ Potenciais relacionados a eventos.

atual é possível observar que existem padrões de ERP encontrados a partir da verificação de violações de caráter sintático e semântico.

Assim, no experimento de Harris et al (2000), nas frases com violação sintática, havia a comparação da anáfora sintática concordando/discordando e, nas frases com violação não sintática, havia a comparação da anáfora logofórica concordando/discordando. A previsão do autor era de que, nas frases que continham uma violação sintática, seria encontrado um P600 por ser de caráter sintático e provocar o padrão semelhante ao encontrado nesses tipos de violação. Já nas frases logofóricas, não seriam encontrados P600 por essa violação ser de natureza semântico-pragmática. Ele ainda afirma que seria difícil saber que tipo de onda se esperar das violações logofóricas por elas não serem definidas na literatura em geral. Esperava-se apenas distingui-las das de padrão sintático. Para tanto, foi feito um estudo com 24 conjuntos de frases experimentais construídos no seguinte modelo:

- (30) a. *The boys' cousin introduced Suzie and himself at the wedding.*
b. *The boys' cousin introduced Suzie and themselves at the wedding.*
c. *The boys' cousin introduced himself at the wedding.*
d. *The boys' cousin introduced themselves at the wedding.*¹⁸

A tarefa feita pelos 40 voluntários consistia em ler as frases e determinar quem recebeu a ação da sentença com base na concordância de número. Apareciam na tela as duas opções de resposta e todos os sujeitos foram incentivados a anteciparem a pergunta e a resposta.

Os resultados confirmaram que as comparações envolvendo restrições sintáticas tiveram um P600 e, nas restrições com logofóricas, não. Sugeriu-se que as comparações envolvendo argumentos são mediadas por processos sensíveis às restrições sintáticas enquanto que para os logofóricas são insensíveis ou sensíveis de forma diferente a essas restrições.

¹⁸ O primo dos meninos apresentou Suzie e ele mesmo no casamento.
O primo dos meninos apresentou Suzie e eles mesmos no casamento.
O primo dos meninos apresentou ele mesmo no casamento.
O primo dos meninos apresentou eles mesmos no casamento.

Ainda buscando compreender o processamento de sentenças em tempo real, tendo como foco a reflexividade e tendo por meio a discussão sobre os campos da sintaxe, de um lado, e da sintaxe + discurso do outro, Petra Burkhardt (2002) faz um estudo em língua inglesa utilizando o paradigma *cross-modal* para verificar o processamento em sentenças contendo logóforos. Ou seja, ambas as posturas foram examinadas na perspectiva do processamento.

Como a interpretação de reflexivos logofóricos envolve acesso a operações não sintáticas, como o acesso às noções de discurso e de ponto de vista, e os reflexivos coargumentos envolvem apenas operações sintáticas, a autora levanta a hipótese de que os contrastes entre esses dois reflexivos teriam custos de processamento diferentes.

Na posição de apenas sintaxe, os dois reflexivos propostos no estudo teriam um comportamento semelhante; já na posição sintaxe + discurso, haveria um contraste entre eles.

Foram feitas as tarefas de compreensão de sentenças e as de decisão lexical, que consistia na apresentação da frase a qual o sujeito ouvia e em seguida respondia à pergunta de compreensão sobre a frase em pontos aleatórios no curso do experimento. Em um determinado momento durante a apresentação da frase, aparecia uma palavra sonda na qual o sujeito teria que decidir apertando o botão “sim” ou “não” se a sonda era ou não uma palavra do inglês. O tempo de reação para a decisão lexical foi gravado.

Foram criados 25 pares de sentenças experimentais e 119 frases distratoras. Nas frases experimentais, cada par consistia em uma frase com um reflexivo coargumento e um reflexivo logofórico, como abaixo:

- (31) a. *The woman_i who was arrogant praised^{PROBE} herself_i^{PROBE} because the network had called about negotiations for a leading role.*
 b. *The girl_i sprayed bug repellent around^{PROBE} herself_i^{PROBE} because there were many mosquitoes in the Everglades.¹⁹*

¹⁹ A mulher que era arrogante elogiou ela mesma porque os contatos tinham ligado a respeito das negociações para um papel de liderança.

A menina pulverizou o repelente em torno dela mesma porque havia muitos mosquitos nos Everglades.

Nas frases com reflexivos logofóricos foram incluídos objetos diretos (*bug repellent*) e preposições de lugar (*around* ou *behind*). Os verbos de ambas as frases foram controlados por frequência e o comprimento total das frases e a distância entre antecedente e reflexivo foram pareados.

As sondas estavam em duas posições: antes do reflexivo (posição controle) e depois do reflexivo (posição experimental). Elas não eram semanticamente relacionadas à sentença nem permitiam que se gerasse uma continuação da sentença. Na posição experimental, é possível medir os recursos necessários durante o processamento do reflexivo, observando se há um custo adicional ou não.

É por meio da comparação dos tempos de reação para a decisão lexical nas duas condições que se pode perceber se há indício para a posição apenas da sintaxe (não há diferença entre as duas condições) ou para a posição sintaxe + discurso (há diferença entre as duas condições com um aumento de tempo nas condições com reflexivos logofóricos).

Os resultados mostram que não houve uma diferença significativa para a sonda na posição controle, mas houve uma diferença significativa para os tempos de reação da sonda (RT) em posição experimental, que registrou um tempo maior para os reflexivos logofóricos, indicando que a interpretação dos reflexivos logofóricos custa mais ao processador do que a dos reflexivos coargumentos.

Os dados obtidos vão a favor de uma abordagem sintaxe + discurso para a reflexividade, sugerindo que a logoforicidade vai além da sintaxe. A diferença entre os reflexivos coargumentos (cuja interpretação é puramente de reflexivos sintáticos) e logofóricos (o que implica que a sua interpretação requer acesso à informação sintática e não sintática) aponta que há um custo adicional para o processador. Esse custo, por sua vez, só é compatível com a posição da sintaxe + posição discurso que afirma que a interpretação de reflexivos logofóricos requer acesso à informação discursiva (extrassintática).

Esses resultados encontrados por Pietra Burkhardt (2002) corroboram os achados de ERP de Harris *et al.* (2000) que descrevemos anteriormente, nos quais os “erros” que envolvem os dois tipos de reflexivos produzem diferentes padrões de ativação de potenciais cerebrais relacionados a eventos, sugerindo a existência de processos sintáticos e extrassintáticos.

Os experimentos feitos em inglês descritos acima são de suma importância para o estudo dos logóforos, pois os fatores pragmáticos têm recebido muita atenção na literatura teórica. Entretanto, não foram encontrados estudos quantitativos determinando a aceitabilidade de sentenças com logóforos. Gibson e Fedorenko (2013) mencionam que a tarefa de julgamento de aceitabilidade é a forma ideal para sondagem do conhecimento sintático/semântico, em comparação com outras tarefas ou medidas de leitura.

Os experimentos em Psicolinguística Experimental com foco no processamento de logóforos têm como base, em sua maioria, a perspectiva de Reinhart e Reuland (1993), em que a posição estrutural da anáfora determina seu caráter logofórico. No presente trabalho, também tomaremos como base a perspectiva de Reinhart e Reuland (1993), visto que observaremos a logoforicidade a partir da posição estrutural da anáfora e do pronome.

4 O experimento

Diante das mais diversas abordagens teóricas para a logoforicidade em várias línguas, o experimento que descreveremos a seguir objetivou verificar a aceitabilidade de construções com anáfora logofórica em PB com a finalidade de compreender se elas estão presentes na gramática de falantes dessa língua, visto que a abordagem de Reinhart e Reuland (1993) mostra que essas estruturas em que a semântica e o discurso operam e exercem uma influência determinante na resolução correferencial da anáfora logofórica são gramaticais em inglês. Além disso, como apresentado anteriormente, esses autores ainda afirmam que, nos contextos em que há uma anáfora logofórica, também pode haver um pronome logofórico.

O estudo de Galves (2001, p. 132 *apud* BRITO, 2009, p. 56) mostra que, em alguns dialetos do PB, o “ele” pode ter leitura anafórica. Consideramos, então, a partir de Reinhart e Reuland (1993), a propriedade logofórica como oposta à propriedade anafórica, em que a propriedade logofórica se caracteriza por ter um antecedente fora do domínio de ligação, seja presente na sentença ou não. Portanto, denominamos o “ele” no experimento como pronome logofórico por este ter um antecedente fora do domínio de ligação nas frases utilizadas. A partir disso, também

nos propomos a verificar a aceitabilidade do pronome logofórico a fim de fazer um contraponto com a aceitabilidade da anáfora logofórica.

A partir dessas abordagens, buscamos responder às seguintes perguntas:

1. Assim como no inglês, essas construções são aceitáveis em português?
2. Há a mesma aceitabilidade em construções logofóricas com anáforas e pronomes?

Para tanto, foi realizado um experimento *off-line* de julgamento de aceitabilidade. De acordo com Leitão (2011):

As aferições obtidas a partir de experimentos *offline* dão informação a respeito da interpretação (momento de reflexão) das frases ou enunciados, ou seja, conseguem capturar reações a estímulos linguísticos quando já houve uma integração entre todos os níveis linguísticos (fonológico, morfológico, lexical, sintático e semântico). (LEITÃO, 2011, p. 223).

Isso posto, o experimento buscou perceber se há aceitabilidade em construções logofóricas em PB de acordo com os estudos de Reinhart e Reuland (1993). Essa aceitabilidade pode ser confirmada através da medida *off-line*, já que os sujeitos emitirão um julgamento para as frases lidas. Assim, utilizamos a anáfora logofórica (AL) *ele mesmo* e o pronome logofórico (PL) *ele* nessas construções para verificar tal aceitabilidade. A hipótese prevista para esse teste foi de que ambas as estruturas seriam consideradas aceitas pelos sujeitos falantes dessa língua, como previsto na Teoria da Reflexividade de Reinhart e Reuland (1993).

4.1 Método

4.1.1 Participantes

Participaram desta pesquisa 77 estudantes do curso de Letras da Universidade Federal da Paraíba (UFPB) – Campus I João Pessoa, todos falantes nativos do português brasileiro, com idade entre 19 e 32 anos.

4.1.2 Material

Para a realização desse experimento, foram criadas 12 frases experimentais, divididas em 2 conjuntos. Cada conjunto continha seis frases com anáfora logofórica *ele mesmo* e seis frases com pronome logofórico *ele*. Além dessas frases experimentais, foram inseridas 24 frases distratoras, sendo 12 frases aceitáveis e 12 inaceitáveis. O propósito dessas frases era fazer com que o sujeito não percebesse qual a natureza das frases que foram o objeto de estudo da pesquisa. Também foi elaborada uma pequena instrução para que o sujeito compreendesse a tarefa. Na tabela abaixo, podemos ver os exemplos das frases experimentais e a instrução contida no teste.

TABELA 2 – Exemplo das sentenças experimentais do teste piloto

INSTRUÇÃO	
Observe as frases abaixo e responda se você as considera como aceitável ou inaceitável. Não há respostas certas ou erradas. Nós queremos apenas verificar suas intuições sobre as frases.	
TIPO DE SENTENÇA	SENTENÇAS
Anáfora logofórica	O irmão de Vítor penteou Lucas e ele mesmo com a escova da sua tia. O genro de Mauro cortou José e ele mesmo com a faca na cozinha.
Pronome logofórico	O irmão de Vítor penteou Lucas e ele com a escova da sua tia. O genro de Mauro cortou José e ele com a faca na cozinha. ²⁰

Fonte: Elaboração própria.

A variável dependente do experimento foi o tipo de julgamento (aceitável x inaceitável) feito em cada condição, e as variáveis independentes foram o tipo de logóforo (anáfora e pronome). A partir disso, obtivemos duas condições experimentais: retomada com pronome logofórico e retomada com anáfora logofórica. Foram controlados o gênero do antecedente e o da retomada, que foram masculinos, e todos os logóforos estavam dentro de uma coordenação.

²⁰ As sentenças experimentais aqui utilizadas seguiram o modelo apresentado por Reinhart e Reuland (1993).

4.1.3 Procedimento

Escolhemos a técnica experimental de julgamento de aceitabilidade (*off-line*), que consiste em averiguar se um dado estímulo é aceitável para os falantes daquela língua, pois a intuição do linguista é insuficiente para atribuir tal julgamento. Essa técnica nos parece ser ideal para testar esse fenômeno devido ao fato de esse julgamento nos proporcionar um dado que torna possível a realização de inferências sobre a representação da logoforicidade na gramática desses falantes.

O experimento piloto foi elaborado por meio do *Google Docs* em um computador e apresentado aos sujeitos por meio do *link* de acesso ao teste. Nesse mesmo *link*, estavam as instruções para a realização do experimento, que apareciam na parte superior do arquivo. Os participantes leram as frases em frente à tela do computador e ao teclado. Ao ler cada frase, o participante teria que escolher entre as opções aceitável e inaceitável segundo seu julgamento. As opções estavam ao lado de cada frase com a possibilidade de apenas uma marcação. As frases foram divididas em dois conjuntos – uma sentença com a anáfora logofórica “ele mesmo” e outra com pronome logofórico “ele”. Dessa forma, o sujeito tinha como tarefa a leitura de cada frase e a escolha de uma opção, ou seja, deveria marcar como “aceitável” ou “não aceitável” cada frase. Os estímulos apresentados no experimento foram randomizados. Como dito acima, nosso intuito era verificar qual sentença seria aceitável e qual seria inaceitável.

4.2 Resultados e discussão

Os resultados obtidos nesse primeiro teste não corroboraram completamente a hipótese inicialmente formulada. Em geral, as sentenças consideradas aceitáveis foram aquelas que continham a anáfora logofórica e não aceitáveis aquelas que tinham o pronome logofórico. A anáfora logofórica teve 276 julgamentos aceitáveis e o pronome logofórico teve 213 julgamentos aceitáveis, porém, neste último, não obtivemos resultado significativo, contrariando a hipótese de que as duas estruturas são gramaticais no PB.

Os resultados dos julgamentos feitos em cada condição podem ser expressos de acordo com a tabela abaixo:

TABELA 3 – Resultados dos julgamentos de aceitabilidade em cada condição experimental

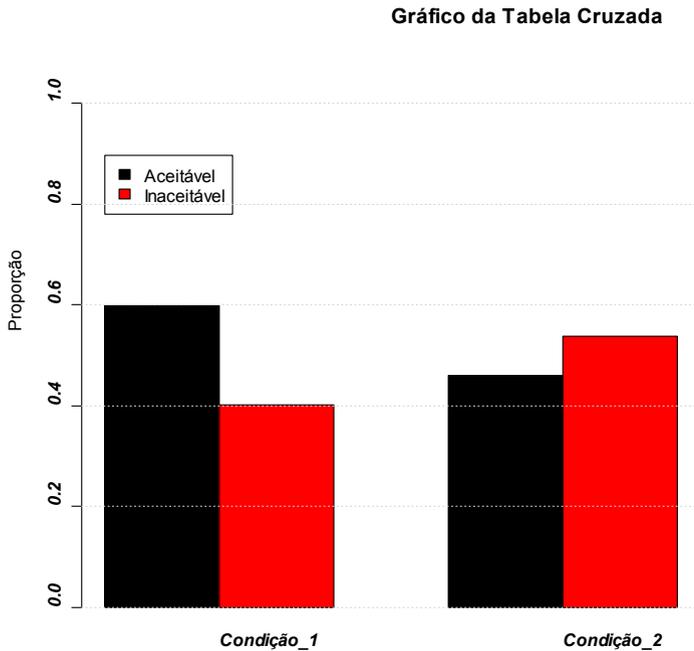
	Condição 1 (retomada com anáfora logofórica)	Condição 2 (retomada com pronome logofórico)	Total
Aceitável	276	213	489
Inaceitável	186	249	435
Total	462	462	924

Fonte: Elaboração própria.

A partir da tabela acima, é possível perceber que houve um maior número de aceitabilidade para a anáfora logofórica (condição 1) e maior número de inaceitabilidade para o pronome logofórico (condição 2).

A diferença da primeira condição, em que obtivemos 276 respostas aceitáveis e 186 respostas não aceitáveis, foi considerada significativa estatisticamente no teste do Qui quadrado de proporção: $\chi^2(1, 462) = 17.532, p < .05$. Essa diferença aponta para uma aceitabilidade desse tipo de construção em português brasileiro. Na segunda condição, com 213 respostas aceitáveis e 249 respostas não aceitáveis, não tivemos um resultado significativo utilizando o mesmo teste estatístico: $\chi^2(1, 462) = 2.805, p = .09$. Este resultado não nos permite apontar que os pronomes logofóricos são inaceitáveis nessa língua. Entre as condições experimentais utilizadas, foi o Qui quadrado de homogeneidade que nos apresentou um resultado significativo ($p < .05$), indicando que os sujeitos viram a anáfora logofórica e o pronome logofórico como estruturas diferentes. O gráfico abaixo expressa esse resultado, mostrando uma tendência inversa na condição PL em comparação com a condição AL:

FIGURA 1 – Gráfico com os resultados do julgamento de aceitabilidade



Fonte: Elaboração própria.

Esses resultados são capazes de nos apontar a necessidade de se fazer outro experimento, a fim de explorar com mais afincado o estudo dos logóforos em PB, visto que esses dados sugerem previamente que a gramática licencia as anáforas logofóricas e as vê como diferentes dos pronomes logofóricos, diferentemente da língua inglesa (REINHART; REULAND, 1993). É possível que o motivo disso seja o *design* do teste não ter contribuído para a aceitação das sentenças com pronome logofórico, pois pode ter acontecido enviesamento, já que os sujeitos que leram as frases com anáfora logofórica também leram as frases com pronome logofórico. Sendo assim, para excluir essa possibilidade, o próximo passo desta pesquisa é realizar um experimento de julgamento de aceitabilidade com o *design* estruturado para que as condições sejam vistas por sujeitos diferentes.

5 Discussão geral

Nossa hipótese no presente estudo foi saber se as construções com anáfora logofórica e pronome logofórico são aceitas pelos falantes do PB. Apesar de nossa intuição poder ser utilizada para considerar tal aceitabilidade, a realidade a ser verificada através dos dados poderia ser oposta, o que nos levou a realizar o experimento descrito anteriormente. Como mostra Gibson e Fedorenko (2013), a intuição do linguista não é suficiente para averiguar tal aceitabilidade.

O estudo aqui reportado aponta para a fiabilidade dos testes de julgamento de sentenças e sua importância no estudo dessas estruturas, posto que não permite que o pesquisador assuma a aceitabilidade de estruturas que são vistas como não gramaticais por sujeitos ingênuos. Logo, este tipo de teste vai além do mero uso da intuição do pesquisador para investigar a gramática do PB. Afinal, se não houvesse esse estudo de aceitabilidade, seria assumido a aceitabilidade das duas estruturas o que não foi sustentada pelo teste de aceitabilidade. No experimento descrito vimos que os sujeitos consideraram aceitáveis as sentenças com anáfora logofórica e não aceitáveis com pronome logofórico.

Os dados obtidos a partir do julgamento de aceitabilidade, feito nesta pesquisa, demonstraram que as estruturas das frases com anáfora logofórica são consideradas aceitáveis pelos falantes nativos do PB. Vamos considerar, assim como Reinhart e Reuland (1993), que esse tipo de estrutura é gramatical. Os dados encontrados apontam para essa possibilidade, já que obtivemos valores significativos para a aceitabilidade dessas estruturas logofóricas.

Na condição experimental anáfora logofórica, os sujeitos julgaram ser aceitável; já na condição pronome logofórico, não houve significância.

Esses resultados são capazes de nos mostrar um caminho a ser explorado no estudo dos logóforos em PB, visto que a gramática licencia as anáforas logofóricas e as vê como diferentes dos pronomes logofóricos, diferentemente de línguas como o inglês (REINHART; REULAND, 1993).

Uma possibilidade para essa diferença é a ocorrência de um enviesamento, pois os sujeitos que leram as frases com anáfora logofórica também leram as frases com pronome logofórico. Sendo assim, para excluir essa possibilidade, faz-se necessário realizar mais um teste de

aceitabilidade com duas condições: anáfora e pronome logofóricos, seguindo o mesmo modelo do teste realizado, porém os sujeitos serão expostos a apenas uma condição experimental.

Por fim, destacamos que a Teoria da Reflexividade prevê que as estruturas logofóricas são completamente aceitáveis pelos falantes do inglês. Assim, a partir da técnica experimental utilizada, concluímos que os dados aqui analisados mostram que a anáfora logofórica, é aceitável. A partir da percepção dos sujeitos participantes do experimento, constatamos que tal estrutura provavelmente compõe a gramática dos falantes do PB.

6 Considerações finais

Os estudos em Linguística, especialmente em processamento linguístico, que investigam o processamento anafórico veem como relevantes os princípios de ligação da Teoria da Ligação na resolução da correferência. Esses estudos levam em consideração a atuação dessas restrições de ligação no momento do processamento, problematizando a gramaticalidade dos antecedentes na resolução correferencial.

No estudo aqui descrito buscamos compreender estruturas com a anáfora logofórica que não são contempladas nessa teoria e compará-las com o pronome logofórico. Para isso, recorreremos à Teoria da Reflexividade de Reinhart e Reuland (1993) como base teórica para o experimento realizado. Procuramos comprovar a hipótese de que frases com logóforos são aceitáveis em PB, já que, segundo Reinhart e Reuland (1993), devem ser consideradas gramaticais em línguas como o inglês, por exemplo.

Com os resultados do experimento, obtivemos evidências de que a anáfora logofórica *ele mesmo* é aceitável em PB ($p < .05$), porém para o pronome logofórico *ele* não obtivemos um resultado significativo ($p = .09$).

Sendo este um dos primeiros estudos referentes à logoforicidade em PB, há um caminho extenso a ser explorado. Este tema ainda necessita de uma abordagem mais ampla tanto no âmbito da Linguística quanto no âmbito da Psicolinguística, visto que neste trabalho estudamos apenas a aceitabilidade de duas entre as diversas estruturas com logóforos. Portanto, além de investigar sua aceitabilidade, é necessário também compreendermos como acontece o processamento dessas estruturas e

investigar experimentalmente o processamento da anáfora sintática e da anáfora logofórica com a finalidade de perceber se essas estruturas diferem em termos de processamento. É relevante mencionar que se faz necessária uma nova investigação com a finalidade de saber como se dá o processamento desses fatores semântico-discursivos nas estruturas com logóforos, bem como observar se o processamento de um elemento logofórico acarretará um maior custo operacional.

Assim, nos próximos passos desta pesquisa, observaremos, em tempo real, a distinção de representação das anáforas sintática e logofórica.

Referências

- ARIEL, M. *Accessing Noun Phrase antecedents*. London: Routledge, 1990.
- BADECKER, W.; STRAUB, K. The Processing Role of Structural Constraints on the Interpretation of Pronouns and Anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, American Psychological Association, v. 28, n. 4, p. 748-769, 2002. <https://doi.org/10.1037/0278-7393.28.4.748>.
- BRITO, D. B. S. *O se reflexivo no português brasileiro*. 2009. Tese (Doutorado em Linguística) – Faculdade de Letras, UFAL, Maceió, 2009.
- BURKHARDT, P. Logophors: Looking Outside of Syntax; Evidence from Real-Time sentence Comprehension. In: THRIFT, E. *et al.* (Ed.). *Proceedings of Console IX, 2002*. Disponível em: <<http://www.hum2.leidenuniv.nl/pdf/lucl/sole/console9/console9-burkhard.pdf>>. Acesso em: 10 abr. 2015.
- CANTRALL, W. R. *On the nature of the reflexive in English*. 1969. Dissertação (PhD) - Universidade de Illinois, Urbana-Champaign, 1969.
- CANTRALL, W. *Viewpoint, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton, 1974.
- CHOMSKY, N. *Lectures on government and binding*. Dordrecht: Foris, 1981.
- CHOMSKY, N. *Knowledge of language*. New York: Praeger, 1986.

CLEMENTS, G. N. The logophoric pronoun in Ewe: its role in discourse. *Journal of West African Linguistics*, West African Linguistics Society, v. 10, p. 141-177, 1975.

FORAKER, S. The processing of logophoric reflexives shows discourse and locality constraints. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, n. 2, p. 295-309, 2003.

GALVES, C. C. *Ensaio sobre gramática do português*. Campinas, SP: Editora UNICAMP, 2001.

GIBSON, E.; FEDORENKO, E. The need for quantitative methods in syntax and semantics research. *Linguagem and Cognitive Processes*, Taylor & Francis Online, v. 28, p. 88-124, 2013.

HARRIS, T.; WEXLER, K.; HOLCOMB, P. J. An ERP investigation of binding and coreference. *Brain and Language*, Elsevier, v. 75, p. 313-346, 2000. <https://doi.org/10.1006/brln.2000.2318>.

KENNISON, S. Comprehending the pronouns her, him, and his: implications for theories of referential processing. *Journal of Memory and Language*, Elsevier, v. 49, p. 335-352, 2003. [https://doi.org/10.1016/S0749-596X\(03\)00071-8](https://doi.org/10.1016/S0749-596X(03)00071-8).

KUNO, S. Functional sentence perspective. A case study from Japanese and English. *Linguistic Inquiry*, MIT Press, v. 3, p. 269-320, 1987.

KUNO, S. Three perspectives in the functional approach to syntax. In: Grossman, Robin E.; JAMES SAN, L.; VANCE, Timothy J. (Ed.). *Papers from the parasession on functionalism*. Chicago: Chicago Linguistic Society, 1975. p. 276-336.

KURODA, S. Where epistemology, grammar and style meet: A case study from Japanese. In: ANDERSON, S.; KIPARSKY, P. (Ed.). *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston, 1973.

LEITÃO, M. M. Psicolinguística experimental: focalizando o processamento da linguagem. In: MARTELOTTA, M. E. *Manual de Linguística*. São Paulo: Editora Contexto, 2011.

LEITÃO, M. M.; PEIXOTO, P.; SANTOS, S. Processamento da co-referência intra-sentencial em português brasileiro. *Veredas on-line*, Juiz de Fora, UFJF, v. 2, p. 50-61, 2008.

MENUZZI, S. *Binding Theory and Pronominal Anaphora in Brazilian Portuguese*. 1997. Tese (Doutorado) - Universidade de Leiden, 1997.

NICOL, J.; SWINNEY, D. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, Springer Link, v. 18, n. 1, p. 5-19, 1989.

POLLARD, C.; SAG, I. A. Anaphors in English and scope of Binding Theory. *Linguistic Inquiry*, MIT Press, 23:2, p. 261 – 303, 1992.

REINHART, T.; REULAND, E. Reflexivity. *Linguistic Inquiry*, MIT Press, v. 24, n. 4, p. 657-720, 1993.

ROSS, J. R. On declarative sentences. In: JACOBS, R. A.; ROSENBAUM, P. S. (Ed.). *Readings in English transformational grammar*. Washington: Georgetown University Press, 1970. p. 222-272.

SELLS, P. Aspects of Logophoricity. *Linguistic Inquiry*, MIT Press, v. 18, n. 3, p. 445-479, 1987.

STURT, P. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, Elsevier, v. 48, p. 542-562, 2003. [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3).

WARSHAWSKY, F. *Reflexivization I & II*. Indiana University Linguistics Club, 1965. [Reproduzido em notas de linguistic underground (*Syntax and Semantics*, v. 7), editado por James D. McCawley. New York: Academic Press, 1976, p. 63-84].

ZRIBI-HERTZ, A. Anaphor binding and narrative point of view: English reflexive pronouns in sentence and discourse. *Language*, Washington, v. 65, n. 4, p. 695-727, 1989. <https://doi.org/10.2307/414931>.

ANEXO 1

As condições experimentais e as frases utilizadas no teste de julgamento de aceitabilidade são apresentadas abaixo:

Retomada com anáfora logofórica
<p>O irmão de André machucou Davi e ele mesmo na BR230. O genro de Mauro cortou José e ele mesmo com a faca na cozinha. O primo de Leandro barbeou Arthur e ele mesmo no banheiro da rodoviária. O irmão de Vítor penteou Lucas e ele mesmo com a escova da sua tia. O amigo de Joaquim feriu Tiago e ele mesmo com o brinquedo pontiagudo. O neto de Breno olhou Murilo e ele mesmo no espelho da sala.</p>
Retomada com pronome logofórico
<p>O irmão de André machucou Davi e ele na BR230. O genro de Mauro cortou José e ele com a faca na cozinha. O primo de Leandro barbeou Arthur e ele no banheiro da rodoviária. O irmão de Vítor penteou Lucas e ele com a escova da sua tia. O amigo de Joaquim feriu Tiago e ele com o brinquedo pontiagudo. O neto de Breno olhou Murilo e ele no espelho da sala.</p>

The acceptability of logophoric anaphor in Brazilian Portuguese

A aceitabilidade da anáfora logofórica em português brasileiro

Flávia Gonçalves Calaça de Souza

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil
flavia.ufpb@hotmail.com

Rosana Costa de Oliveira

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil
rosana.ufpb@gmail.com

Judithe Genuíno Henrique

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil
judithegh@gmail.com

Abstract: This study sought to investigate whether or not these logophoric structures are acceptable or not for the speakers of Brazilian Portuguese (BP). Both syntactic and semantic approaches have been used to explain logophoricity and a strong theoretical discussion has been conducted as to what kind of approach would be best to account for this phenomenon. Authors such as Reinhart and Reuland (1993) points out that logophoric anaphora exists separately from syntactic anaphora in Universal Grammar. The syntactic anaphora is guided by syntactic factors, such as location and c-command. By contrast, logophors may or may not notice these syntactic conditions. These studies also demonstrate that, in cases of logophoricity, the pronoun and anaphora are interchangeable, therefore, either one or the other can be used in the same position in the sentence. Based on these theoretical assumptions, we examine the acceptability of logophoric buildings in BP in order to know if they are present in the grammar of the speakers of that language.

An offline experiment of acceptability judgment was conducted, given that, according to Gibson and Fedorenko (2013), linguistic intuition is insufficient to verify such an acceptability. The following experiment demonstrated that the results were significant for logophoric anaphoras, which shows that a degree I anaphoras and logophoric pronouns are different. No significant result was observed for the pronoun, which does not allow us to point out that are not licensed in BP.

Keywords: judgment of acceptability; logophoricity; anaphora; pronoun.

Resumo: Este trabalho buscou investigar se as estruturas logofóricas são aceitáveis ou não para os falantes do português brasileiro (PB). Diversas abordagens sintáticas e semânticas têm sido usadas para explicar a logoforicidade, e uma forte discussão teórica é sobre qual tipo de abordagem melhor dá conta do fenômeno. Autores como Reinhart e Reuland (1993) apontam que a anáfora logofórica existe separadamente das anáforas sintáticas na Gramática Universal. As anáforas sintáticas são guiadas por fatores sintáticos como localidade e c-comando. Já a logófora pode ou não observar essas condições sintáticas. Esses estudos também demonstram que, nos casos em que há logoforicidade, a anáfora e o pronome são intercambiáveis, ou seja, tanto um quanto o outro podem ser usados na mesma posição na frase. Com base nesses pressupostos teóricos, examinamos a aceitabilidade de construções logofóricas em PB com a finalidade de saber se elas estão presentes na gramática dos falantes dessa língua. Foi realizado um experimento *off-line* de julgamento de aceitabilidade, visto que, segundo Gibson e Fedorenko (2013), a intuição do linguista não é suficiente para averiguar tal aceitabilidade. O presente experimento mostrou que a anáfora logofórica é licenciada em PB, o que aponta que os sujeitos veem a anáfora logofórica e o pronome logofórico como diferentes. Não houve resultado significativo para o pronome, o que não nos permite apontar que não são licenciados em PB.

Palavras-chave: julgamento de aceitabilidade; logoforicidade; anáfora; pronome.

Received on: December 10, 2016

Approved on: April 04, 2017

1 Introduction

The Binding Theory (CHOMSKY, 1981), seeks to study the coreference relations that exist between nominal and pronominal expressions, within the scope of a sentence. To explain such relations, Chomsky (1981) proposes three principles that govern these relations, namely:

- (1) Principle A – An anaphor must be A-bound in its binding domain.¹
- Principle B – A pronominal must be A-free in its binding domain.
- Principle C – An R-expression must be A-free.²

Principle A, which refers to reflexive and reciprocal anaphors (self, one another), predicts that an anaphor must be bound in its binding domain, that is, be bound by an antecedent that is c-commanded in its local domain. A c-command is understood as the syntactic notion of binding between two constituents. Let us observe the example below:

- (2) [*O neto do João*]_k³ *se*_{i/*k} *adora*. ([João's grandson]_k loves himself_{i/*k}.)

In this example, the DP⁴ “*O neto do João*” (João’s grandson) is the antecedent of the anaphor “*se*” (himself); therefore, the DP “*O neto do João*” (João’s grandson) c-commands the anaphor “*se*” (himself).

Principle B refers to the pronouns, and states that all pronouns are free in their binding domain, that is, the antecedent and the pronoun may not be on the same sentence. Let us see the examples below:

¹ A binding domain is understood as being the limit presented by the anaphora and its antecedent..

² Binding Principles (CHOMSKY, 1981)

A. An anaphor must be A-bound in its binding domain.

B. A pronominal must be A-free in its binding domain.

C. An R-expression must be A-free.

³ Indexation index..

⁴ Determiner Phase

- (3) a. [*Penha_i acha que [Bruna vestiu ela_i com o vestido.]*]
 ([Penha_i thinks that [Bruna dressed her_i with the dress.]])
 b. **[Penha acha que [Bruna_i vestiu ela_i com o vestido.]*]
 ([Penha thinks that [Bruna_i dressed her_i with the dress.]])

Sentence (3a) is absolutely plausible, given that the antecedent is outside the pronoun domain, unlike sentence (3b), which is agrammatical, since we have the pronoun and its antecedent in the same domain, as contended by the Binding Theory.

Finally, there is Principle C. This principle mentions referential expressions, which must be free in all occurrence contexts, since they have referential autonomy and, therefore, do not need to be bound. Let us observe the example below:

- (4) *O João observou o Pedro na sala.* (João observed Pedro in the room.)

In some cases, the anaphor is not constrained by the binding conditions proposed by Chomsky (1981), as shown by Reinhart and Reuland (1993). These authors refer to this type of anaphor as logophoric. As we can see:

- (5) *Maria contava cinco pessoas na cozinha, além de si mesma/dela.*
 (Maria counted five people in the kitchen, in addition to herself/her.)
 (6) *Carlos viu uma blusa perto de si/dele.* (Carlos saw a shirt next to himself/him.)

In the examples above, the anaphor is in an adjunct prepositional phrase (PP)⁵ and is not an argument required by the verb. According to Reinhart & Reuland (1993), Chomsky (1981; 1986) does not explain the contexts exemplified in (5) and (6), precisely because the anaphor and its antecedent are not coarguments. These anaphors do not behave as predicted by the Binding Theory, since it does not explain the coreference with antecedents outside the binding domain.

⁵ Prepositional phrase

Based on these and other examples, Reinhart and Reuland (1993) observe the anaphors from the argument standpoint, indicating that the anaphor is logophoric when it is not in an argumental position. The syntactic anaphor, as its name indicates, is guided by syntactic factors such as location and c-command, while the logophoric anaphor is related to discursive factors. This is shown by the examples below:

- (7) *João se vangloriava de que o prefeito convidou Ana e ele mesmo para um café da manhã.* (João boasted that the mayor invited Ana and himself to have breakfast.)
- (8) *Vânia encontrou tempo para verificar que, além dela mesma, havia um outro funcionário que tinha sido demitido.* (Vânia found time to verify that, besides herself, another employee had been fired.)

The sentences above reveal that Chomsky's (1981; 1986) predictions for anaphors may not be verified, since the anaphor is not in the binding domain in either of the sentences above. Zribi-Hertz (1989) provides approximately 130 examples, such as those provided above, of sentences in which that anaphor is not in an argumental position.

Reinhart and Reuland (1993) also state that it is possible to have a logophoric pronoun whenever there is a logophoric anaphor, as in (9) and (10):

- (9) *Carlos estava chateado quando Júlia colocou Bruno e ele mesmo em perigo no paraquedismo.* (Carlos was upset when Júlia put Bruno and himself at risk in parachuting.)
- (10) *Carlos estava chateado quando Júlia colocou Bruno e ele em perigo no paraquedismo.* (Carlos was upset when Júlia put Bruno and him at risk in parachuting.)

The type of structure present in sentences (9) and (10) is discussed at length by both Pollard and Sag (1992) and Reinhart and Reuland (1993), as this is consistent with the fact that, in English, the pronoun is an acceptable means of alternating with the anaphor, maintaining the same interpretation.

The present work focuses on determining the acceptability of the logophoric anaphor and the logophoric pronoun in Brazilian Portuguese (BP), that is, this study's goal was to investigate if subjects consider sentences with logophoric anaphors and logophoric pronouns to be acceptable or unacceptable according to their intuition.

Gibson and Fedorenko (2013) indicate that, although the researcher's intuitions are useful to several studies, many others do not allow for high reliability based only on intuition, and require a stricter methodological control to detect the subtleties and to distinguish among different theoretical positions. The author also states that a major problem in the syntax field of study is that many articles do not include the experimental evidence to support their research hypotheses. The author also recommends that the studies collect quantitative evidence to ensure an improvement in this field of study. Taking such importance into account, this research focuses on the Acceptability Judgment quantitative method, in which the subjects evaluated sentences with logophoric anaphor and logophoric pronoun with the intent to present reliable data on the acceptability of these sentences.

The key goal of the off-line acceptability judgment performed was to observe the acceptability of anaphors deemed as logophoric, especially the anaphor "*ele mesmo*" ("himself") in BP, in order to compare it to the pronoun "*ele*" ("him"), to understand if, similarly to the English language, this phenomenon is present in the grammar of BP speakers. By hypothesis, based on Reinhart and Reuland's (1993) Reflexivity Theory, these theories would be acceptable and would not violate the binding conditions, given that they are not constrained to them.

It is important to note that there is a lack of studies in BP on the resolution of logophoric structures, as well as the processes it requires. Therefore, this work, in addition to verifying the acceptability of logophoric structures in BP, sought to conduct a bibliographic review of the studies performed on this topic in other languages, which indicate that logophoric anaphors occur separately from syntactic anaphors in Universal Grammar (UG).

2 Logophoricity in Linguistics

Although the relations involving anaphors, pronouns, and their antecedents in some cases are established in compliance with the constraints imposed by the Binding Theory (CHOMSKY, 1981), some studies (REINHART; REULAND, 1993; ZRIBI-HERTZ, 1989) show that these forms may be interpreted otherwise. As indicated above, logophoricity occurs in these cases, and it has been characterized and defined based on several perspectives. One of these perspectives is that it is not subject to the c-command and position constraints imposed by the Binding Theory for having a logophoric status.

Chomsky's (1981), Reinhart and Reuland's (1993) and Zribi-Hertz's (1989) approaches will be outlined in detail below, to provide a descriptive survey of this scarcely researched topic. Some of these approaches view the logophoric anaphor from the c-command configuration, that is, not only limiting and defining it by means of constraints of a syntactic nature, but also observing the discourse and semantics prior to the syntax. Zribi-Hertz's (1989) approach, for instance, gives theoretical priority to discursive factors. If these are not sufficient, the syntactic domain is sought.

2.1 Chomsky (1981)

The Binding Theory formulated by Chomsky (1981) describes the syntactic constraints on anaphors, pronouns, and referential expressions. He postulated three binding principles: A, B, and C. Principle A predicts that an anaphor must be bound to the antecedent that is within its domain, that is, it must be bound to the antecedent within the minimum clause that contains the anaphor. In addition to the presence of a local antecedent, principle A also predicts that the antecedents c-command the anaphor. The sentences below are two examples of anaphors bound to their local antecedent:

- (11) *[A prima de Joana]_i se_i acha muito bonita.* ([Joana's cousin]_i finds herself_i very pretty.)
- (12) *O Carlos_i machucou ele mesmo_i enquanto cozinhava.* (Carlos_i hurt himself_i while cooking.)

Examples (13) and (14) below show a DP that is not c-commanding the anaphors *se* (herself) and *ele mesmo* (himself), since, according to the Binding Theory, the anaphor *se* (herself) may not refer to Joana, and the anaphor *ele mesmo* (himself) may not refer to Carlos, making the sentences ungrammatical.

(13) **A prima de [Joana]_i se_i acha muito bonita.* (*[Joana's]_i cousin finds herself_i very pretty.)

(14) **O Carlos_i machucou Vivian e ele mesmo_i enquanto cozinhava.* (*Carlos_i hurt Vivian and himself_i while cooking.)

Principle B predicts that the pronoun is free in its binding domain. In sentence (15) below, the pronoun *ela* (she) may not have the *Maria* DP as its antecedent, since it is within its binding domain.

(15) **A Maria_i adora ela_i.* (*Maria_i adores her_i.)

Principle B, therefore, imposes that the pronoun may not be bound to its local antecedent, differently from principle A, which states that the anaphor must be bound to this antecedent.

Principle C, however, determines that referential expressions must be free in any syntactic context, that is, they may not be bound. In (16), the R-expression *o animal* (the animal) does not have antecedent within the sentence, which makes it ungrammatical. R-expressions do not require an antecedent element to extract their meaning, that is, they are referentially independent, which sets them apart from anaphors and pronouns, which have a referential dependency.

(16) *O animal fez um grande barulho.* (The animal made a great noise.)

Also in accordance with the Binding Theory, anaphors and pronouns are in complementary distribution. That is, an element can only be in the sentence where the other cannot. Hence, for a sentence to be deemed grammatical, these elements must be in opposite environments. An anaphor must be bound and a pronoun must be free. Therefore, if an anaphor has an antecedent within the same clause that includes it, it is grammatical. However, if this very sentence included a pronoun rather than an anaphor, the sentence would be ungrammatical. Nevertheless, several authors have demonstrated that this complementary distribution

may not always be identified, since sentences such as (17) are considered to be grammatical.

- (17) *Mateus_i sabia que ninguém gosta de João e dele mesmo_i/dele_i por causa de suas personalidades.* (Mateus_i knew that no one likes João and himself_i/him_i because of their personalities.)

Reinhart and Reuland's (1993) proposal described below recommends that anaphors, such as the one in example (17), be observed based on the notion of arguments and, if they are not coarguments in the same predicate, they will still be deemed grammatical. The fundamental idea is that complementary distribution does not exist in sentences such as (17), which contains a logophoric anaphor.

2.2 Reinhart and Reuland (1993)

The Reflexivity Theory proposed by Reinhart and Reuland (1993) is based on the need to reformulate the Binding Theory (CHOMSKY, 1981), explained by the many problems posed by principles A and B in this theory. The authors also question the complementary distribution between anaphors and pronouns, identifying several examples in which this complementarity fails. This theory also makes the difference clear when there is a connection and when there is a coreference between the anaphoric forms.

For such, the authors propose returning to the interpretation of reflexive predicates in the natural languages in which the verb has a strict connection with the reflexivization phenomenon. Based on this, reflexivization would then be analyzed as a property of predicates rather than a property of the anaphor, that is, the distribution of anaphoric forms is done based on the semantic properties of anaphoric forms, in this case, the property of reflexivizing a predicate (*inherent properties*). Therefore, the following anaphor definition is assumed:

Anaphors (of both the SE and the SELF type) are referentially defective NPs, which entails, for example, that they cannot be used as demonstratives, referring to some entity in the world (though it does not entail that they must be bound variables). (REINHART; REULAND, 1993, p. 658).

A typology of anaphoric expressions is proposed, based on this definition and on both types of anaphor functions. SE and SELF anaphors are included in the same group regarding the referential dependency property, and SE anaphors and pronouns are grouped because they do not have the reflexivization function. The summary of this typology is presented in the table below:

TABLE 1 – Reinhart and Reuland’s (1993) proposal for anaphor and pronoun distribution

	SELF	SE	PRONOUN
Reflexivizing function	+	-	-
R(eferential Independence)	-	-	+

Source: Reinhart and Reuland (1993)

According to the table above, SELF anaphors have the property of reflexivizing a predicate (+ reflexive function), but they do not have referential independence. The authors consider that this SELF reflexive function is sensitive to the binding condition proposed by them. SE anaphors, however, neither have referential independence, nor are responsible for making a predicate reflexive (- reflexive function), since they are only markers for inherently reflexive predicates. Pronouns, by contrast, do not reflexivize predicates (- reflexive function); however, unlike the anaphors, they do have referential independence, given that they carry person, gender, and number traits – although not necessarily all of these.

For the authors, however, it should be noted that both types of anaphors mentioned can be used logophorically. However, since studies that establish the distinction between the anaphor’s grammatical and logophoric functions are scarce, there are several mistakes regarding the syntactic distribution of anaphors and the terms local and long distance, given that, if used logophorically, both types of anaphors may result in all types of distance. Hence, logophoric reflexives do not follow a specific rule, but they should be considered to be grammatical. Nevertheless, the authors indicate that one of the properties of the logophoric anaphor is not being c-commanded (REINHART; REULAND, 1993, p. 660).

Reinhart and Reuland (1993) discuss that it is possible to have a pronoun whenever there is a logophoric anaphor. By stating this, the authors do not provide an explanation for the reasons why preference is given to a logophoric anaphor or to a pronoun. In a note, the authors themselves suggest Ariel's (1990) approach, which takes the most accessible available candidates in the discourse into account.

Regarding the grammatical function, the authors state that the domains are limited to two: local and long distance. The local domain corresponds to the reflexivity domain in which the SELF anaphor mandatorily reflexivizes the predicate. This domain is regulated by Conditions A and B, reformulated by them, which will be described below. The long distance domain, however, refers to the binding domain of SE anaphors, which adopt the pronoun pattern, and, therefore, are governed by Condition B.

Conditions A and B must be read as requirements and are defined as follows:

- (i) Condition A: a syntactic predicate marked reflexively is reflexive.
- (ii) Condition B: a reflexive semantic predicate is reflexively marked.

Condition A applies to predicates that are marked reflexively by syntactic ways, while Condition B inherently (semantically) applies to reflexives. A syntactically reflexive predicate is understood as being one in which two of its arguments are coindexed (with indistinct gender, number, and person traits), while a reflexively marked predicate is understood as being either lexically reflexive (which is marked by the presence of an SE anaphor) or having a SELF anaphor as one of its indexed arguments. Therefore, it is not the anaphor in itself, but rather the reflexive marking that allows for it to be interpreted reflexively.

Thus, for Reinhart and Reuland, Conditions A and B are not related to the c-command syntactic configuration, that is, with the possibility of binding to the antecedent available on the same pronoun or anaphor domain, but are related to the intrinsic property of the predicate being inherently reflexive or not.

To observe how Conditions A and B apply, let us observe the examples below:

- (18) a. **John_i likes him_i*.
 b. *John_i likes himself_i*.
 c. *John_i said Ann likes him_i*.

It should be noted that in (18a) *John* and *him* are coarguments of the predicate *likes*. This syntactic predicate will not be reflexive because it was not marked by the SELF anaphor corresponding to the anaphor that reflexivizes predicates, but it was instead replaced with a pronoun, making the phrase ungrammatical. A predicate, such as *like*, would be subjected to condition A – a predicate that would be syntactically marked as reflexive, since one of its arguments would be a SELF anaphor. In (18b), the phrase is grammatical, since the SELF anaphor is used in the sentence as a reflexive marker and is compatible with Condition A. In (18c), however, *John* is a coargument for *said*, and *him* is a coargument for *likes*. Hence, the predicate is not reflexive, since there is no coindexation between the arguments in the same predicate; therefore, it is not guided by Condition B. In (18c), the predicates are not semantically reflexive; therefore, they are not reflexively marked.

One of the problems the authors identified about Condition A in Chomsky's Binding Theory (CHOMSKY, 1981) is the occurrence of anaphors that are free of their domain, as demonstrated by Kuno (1987) and Zribi-Hertz (1989) with SELF anaphors in the first, second, and third persons. Let us see:

(19) *Max boasted that the queen invited Lucie and himself for a drink.*

(20) **Max boasted that the queen invited himself for a drink.*

Reinhart and Reuland (1993) mention that, for Zribi-Hertz (1989), long-distance bound, third-person anaphors violate principle A and its use is known as logophoric. Nevertheless, the authors argue that it is not possible to conclude that the contexts in which there is a point of view⁶ allow for a violation of Condition A. For Reinhart and Reuland (1993), it is not merely a discursive issue, but, above all, a structural issue, given that (20) the SELF anaphor is an argument of its

⁶ The notion of point of view has been characterized in the literature to indicate the coreference of the anaphor/pronoun and the self-speaker, as proposed by Kuno (1987), whose discourse, thoughts, and feelings are being reported.

predicate and, therefore, the predicated is marked reflexively. In (19), there is no reflexively marked predicate, as the argument for *invited* is not only *himself* (SELF anaphor), but also *queen*, with which *himself* does not share gender traits; therefore, the first may not be coindexed the latter.

The view developed by the authors, based on the distinction between the grammatical and logophoric functions, show that Condition A is applied only to the anaphor in the argumental position. That is, SELF anaphors that occur outside this position (logophoric) are exempt from this condition.

In this discussion, the authors also state that the distinction between the anaphoric and the logophoric use of SELF is superfluous, given that, syntactically, there is only one type of anaphor, whose use is governed by Condition A. This immediately excludes the SELF anaphor in which there are no reflexive predicates as applicable to this Condition. Therefore, logophoricity is not coded in syntax, and the logophor does not need to be bound/coindexed to an antecedent. Its relation may be that of coreference (REINHART; REULAND, 1993, p. 673). This position is in disagreement with Sells (1987), who states that logophors, especially perspective logophors, are bound.

Only one case has been identified in which Condition A allows for the logophoric use of the SELF anaphor, that is, when it does not have an argumental position and, consequently, does not mark the predicate as reflexive.

As stated above, Condition A governs the anaphor. In addition, the Reflexivity Theory also contends that the pronoun is not excluded in certain contexts in which the prediction of an anaphor occurs. An example of this context is the occurrence of an *NP picture*, as in (21), in which an additional mechanism is provided to justify the occurrence of a pronoun.

(21) *Mary saw a Picture of herself.*

Moreover, it is important to emphasize that one of the properties of logophoric anaphor mentioned by the authors is that it does not need to be c-commanded by its antecedent. Therefore, the *NP picture* examples are easily accepted, as coreference rather than binding occurs.

2.3 Zribi-Hertz (1989)

Zribi-Hertz's article published in 1989 presents a detailed survey of the occurrence of reflexive pronouns in English that are compatible with the Chomskyan Binding Theory, and observes that some of these anaphors violate this theory's principle A, which reveals that it is based on a relatively small data corpus and that it is not sufficient to prove that the distribution of daily use pronouns is as constrained as the prediction made by principle A.

These pronouns that violate such a principle, according to the author, draw a clear line between syntax and discourse, and show that Chomsky's principle A is complete only if it is an internal theory of the phrase, not taking the discourse into account, although it is incomplete, as it ignores an integral component of the reflexives' grammar, and thus does not account for several data sets.

The key discussion in Zribi-Hertz's study is to understand where the line should be drawn, regarding the anaphor, between syntax and discourse. A survey of the syntactic and semantic properties of reflexive pronouns in English is conducted with the intent of confronting them with a corpus of English texts. After the corpus analysis, the author explains that English reflexive pronouns may be long-distance bound and may violate several structural constraints. Based on this, a relation between the grammar of locally bound reflexive and the discursive grammar of long distance bound reflexives has been proposed.

Some of the problematic data that arose after the Binding Theory jeopardize the complementarity between anaphors and pronouns. These include: NPs Picture, genitive positions, some PPs, and emphatic contexts, which are discussed by Warshawsky (1965), Ross (1970), Cantrall (1974), Kuno (1987), among other authors, as can be seen below:

- (22) a. *They thought that [pictures of {them / themselves}] would be on sale.*
 b. *We thought that [John's pictures of {us / ourselves}] would be on sale.*
- (23) *John said that there was a picture of {him / himself} in the post office.*

- (24) a. *(Mary thought that) [a picture of {you / yourself}] would be nice on the wall.*
 b. *(Mary thought that) [a picture of {me / myself}] would be nice on the wall.*
- (25) *They heard the stories about {them / themselves}.*

Using these and other counter-examples, some authors either indicate a change or alteration in the typology to include them within the structural theory, or suggest that these occurrences are outside the syntax field. However, this article argues that a grammatical theory for English reflexive pronouns cannot be complete without a discursive component.

The author quotes Cantrall's (1969) discussion about the hypothesis that, in English, whenever it is possible to alternate pronouns and anaphors in the same structural context, the choice for one or the other expresses the choice of a narrative point of view, as well as quotes Kuroda's (1973) discussion, which suggests that the *zibun* reflexive option is related to the non-reporting narrative style by means of which the author reports the events "from within" (a given character) contrary to the author's own point of view.⁷

Cantrall's (1969) discussion may also be observed in the examples below:

- (26) a. *The women_i were standing in the background, with the children behind them_i.*
 b. *The women_i were standing in the background, with the children behind themselves_i.*

Sentences 26a and 26b do not have the same informative content, since in 26b the children are located "behind the women", from the internal point of view of the discursive protagonist (the women), and, in 26a, the children are "behind the women", from the speaker's point of view. The contrast is related to the point of view option, which may be that of a discourse protagonist or that of the speaker and, consequently, as Cantrall (1969) points out, in choosing a third person in pronominal anaphors and pronouns, there is no structural conditioning.

⁷ The point of view is also the key concept of pronoun analysis made by Kuno (1972, 1983, 1987).

The proposal in Chomsky's Binding Theory does not correlate to the semantic properties above, and it is assumed that they do not belong to the syntax and derive from their structural properties. However, the author, although agreeing that the point of view does not belong to syntax, disagrees with its derivation from structural properties: "*It is the structural properties of pronouns that are, in a sense, derived from their discourse properties*".⁸ (ZRIBI-HERTZ, 1989, p. 705)

To justify this statement, the examples in (27), in which there is a semantic contrast, are provided:

- (27) a. *John_i hid the book behind himself_i.*
 b. *John_i hid the book behind him_i.*

For Zribi-Hertz, the spatial relation between John and the book is understood as being more direct in 27a than it is in 27b, given that, in the latter, John hid the book somewhere behind the shoulder line, whereas in 27a John hid the book very close to himself, probably in contact with his own body.

It is thus argued that the "subject of consciousness" is not a semantic variation of the syntactic subject, but stands out as a linguistic concept belonging to discourse grammar, and that this subject of consciousness, rather than the syntactic subject, is the relevant concept for the grammar of long-distance bound reflexives in English.

The concept of "subject of consciousness" is a category of discourse grammar, being similar to the concept of logophoricity proposed by Kuno (1987) and Clements (1975). As the author states, the "subject of consciousness" is a semantic property attributed to a referent whose thoughts or feelings, optionally explicit in the discourse, are transported by a part of the discourse. This "subject of consciousness", still in the author's words, is generally understood as being [+ human].

The author states that reflexive pronouns belong to discourse grammar, from which sentence grammar is a specific sub-domain. With this conclusion, the author suggests that structural constraints may have their discursive motivation.

⁸ "São as propriedades estruturais dos pronomes que são, em certo sentido, derivadas de suas propriedades discursivas."

An important statement made by the author is that the anaphoric relations indicated in her study belong to discourse grammar, since the structural constraints come into play when the discourse principle is no longer relevant.

Therefore, Zribi-Hertz's proposal argues that any anaphor structural theory must be completed by the discourse principles, since, to account for this phenomenon, the notions of discourse structure must be taken into account.

This proposal of sentence (syntax) grammar being a sub-part of discourse grammar authorizes reflexive pronouns to violate Chomsky's Principle A, in case some discursive properties are met. This makes Zribi-Hertz's proposal different from Reinhart and Reuland's (1993), since, for the latter, the syntax operates primarily and, if it is not sufficient, semantics and discourse are free to act.

3 Logophoricity in Psycholinguistics

Despite the several approaches to the phenomenon, few studies in Experimental Linguistics have understood the processing of this type of structure as their primary focus. In this topic, this study will demonstrate some approaches in the inter-sentence processing and some experiments developed in Psycholinguistics.

In the field of inter-sentence Experimental Psycholinguistics, there is a discussion about the online processing of sentences that have a single clause being constrained exclusively to the action of the Binding Principles proposed by Chomsky (1981) in the first processing stages. This problem includes a strong debate about the Chomsky's binding principle's moment of action, in the course of processing, which has unfolded into a differentiated prediction with the models that argue how much processing is affected by its antecedents, namely, the models: initial filter, reversible filter, and interactive filter.

In general lines, the first model was presented by Nicol and Swinney (1989), who state that the binding between the antecedent and the anaphor is constrained to Chomsky's (1981) binding principles, both in the initial stages of processing and in the subsequent stages. The second model is defended by Sturt (2003), Kennison (2003), and Leitão, Peixoto and Santos (2008) by presenting evidence of a reversible filter in which these principles guide the processing in

the initial stages but may be violated afterwards by many factors. Badecker and Straub (2002), however, analyze their data interpreting that not only the syntactic constraints operate in this first stage, but also other constraints of discursive nature. These last authors propose the interactive model.

The model described by Nicol and Swinney (1989) shows that the constraints proposed by the Binding Theory are applied in the initial and subsequent stages of processing. According to this model, the antecedent for anaphor is immediately chosen according to the binding theory, while other antecedents are disregarded in these first stages of processing and afterwards.

The authors analyzed the processing of pronouns and anaphors that had two types of antecedents, according to the Binding Theory: available and inaccessible. They performed *priming cross modal*⁹ experiment to observe the action of binding constraints. In this experiment, participants heard the sentences and made a lexical decision after viewing a probe word that appeared after anaphors and pronouns.

Therefore, it was predicted that the structurally available antecedents would influence the processing of coreference in the initial stages and the inaccessible antecedents would be immediately excluded and disregarded in the subsequent interpretation (Initial Filter Hypothesis).

The results of the study performed by them brought evidence to the Initial Filter Hypothesis, since there was a significant effect for the structurally available antecedents, while there was no significant effect for the structurally inaccessible antecedents, both for the anaphor and the pronoun. This indicates that they were not taken into account at the time of the coreference resolution and there is a grammatical filter that determines which candidates may refer to the anaphor and the pronoun.

Nevertheless, Sturt (2003), Kennison (2003), and Leitão, Peixoto and Santos (2008) found evidence against the initial binding filter model. Their data show that binding constraints are applied at the initial stages but may be violated afterwards. The filter would thus be reversible, with the possibility of being violated at a second moment of processing due to several factors.

⁹ When stimuli presented in the priming and target are of different modalities.

Sturt's (2003) work has shown that binding constraint are relevant in processing coreference at the first processing stage. Discursive factors would only produce late influence, which justifies the effect of inaccessible antecedents found in their study.

Running in line with Sturt's (2003) findings, Kennison (2003) investigated the action of principle B in processing by means of the self-monitored reading, capturing the processing in two stages.

The author conducted an experiment using the pronouns *her*, *his*, and *him* with available and inaccessible antecedents, in which the subjects took longer to read the pronoun under the condition in which the structurally inaccessible antecedent had the same gender in the anaphora, when compared to the condition in which the gender was the same.

Based on the results, Kennison (2003) suggests that the structurally available and inaccessible antecedents are considered during the coreferential resolution, differently from Nico and Swinney's (1989) proposal, in which only structurally accessible antecedents are taken into account.

Leitão, Peixoto and Santos' (2008) findings for BP, identified by means of two self-monitored reading experiments, dialog with Kennison's (2003) and Sturt's (2003) results. In the experiment conducted by the authors, the processing of the pronoun "*ele*" (him) as an object was investigated. At first, no significant difference in reading the critical segment (pronoun) was identified, which suggested that principle B blocked, in the first stage, the possibility of the pronoun "*ele*" (him) to be bound to the subject.

The reading times were longer when the structurally inaccessible subject had the same gender, number and animacy as the pronoun than under the conditions in which these traits did not match. These results suggest that coreference processing occurs in two stages.

A preamble was added to the second experiment with an available antecedent that shared the pronoun traits. The presence of this available antecedent made the pronoun reading significantly longer than it was in experiment 1, that is, there was no influence of inaccessible antecedents.

Finally, Badecker and Straub (2002) propose a model in which several constraints operate at the initial processing moment. Both syntactic and discursive constraints are combined in parallel, that is, the binding constraint act along with other types of information (gender, number and discursive focus traits). In this understanding, both available and inaccessible antecedents are relevant in coreferential processing from the first processing stage.

The self-monitored reading study conducted by them included both structurally accessible and inaccessible antecedents. Badecker and Straub (2002) found results in the post-critical segment, which were read faster when the anaphora gender matched the antecedent gender. However, the reading was slower when the gender of inaccessible and accessible antecedents matched the anaphora gender, which shows that inaccessible antecedents also affected the sentence processing. In this line, the authors propose that inaccessible antecedents interfere in the processing.

However, these models mentioned above do not specifically address the discussion of the processing of logophoric anaphors and logophoric pronouns, although they mention the possibility of certain types of structure in which the anaphor is used and in which the binding principles do not operate.

In another psycholinguistic study, specifically involving logophoricity, Foraker (2003) used the SELF form logophorically to examine if discursive information used in interpreting this logophoric is similar to that used in pronoun interpretation. In this study, a self-monitored reading was performed in which the distance between the logophor/pronoun and its antecedent was controlled, with three possible positions for the antecedent (*First-mentioned, Middle, and Most Recent*). In this experiment, Foraker shows that there was a *Spillover* effect (after the logophor/pronoun), which revealed that logophors and pronouns were processed in a similar manner only under the conditions in which the antecedent was at a long or intermediate distance in relation to the anaphora. When the antecedent was mentioned more recently, the logophor was read more quickly than when it was at a distant or intermediate position.

Their proposed explanation for the results found under the conditions with long and intermediate distance antecedent is that the coreference resolution for the logophoric anaphor appears to use the same type of pragmatic and discursive information used for pronominal coreference, given that, both for logophors and for pronouns, the same processing pattern was found. The set of sentences in the experiment is exemplified in (28) and (29) below:

(28) *Logophoric Reflexive*

First-mentioned: Megan wondered/ if Isaac had found out/ that Rick wanted to invite/ Sally and herself/ to the birthday party.

Middle: Isaac wondered/ if Megan had found out/ that Rick wanted to invite/ Sally and herself/ to the birthday party.

Most Recent: Rick wondered/ if Isaac had found out/ that Megan wanted to invite/ Sally and herself/ to the birthday party.

(29) *Pronoun*

First-mentioned: Albert was upset/ when Debbie didn't care/ that Rachel had endangered/ Gordon and him/ on the climbing trip.

Middle: Debbie was upset/ when Albert didn't care/ that Rachel had endangered/ Gordon and him/ on the climbing trip.

Most Recent: Rachel was upset/ when Debbie didn't care/ that Albert had endangered/ Gordon and him/ on the climbing trip.

Moreover, for the English language, Harris *et al.* (2000) conducted a study using EEG¹⁰ with the intent to determine the characteristics of the ERP¹¹ responses in sentences with syntactic and nonsyntactic violation, since in current literature it is possible to observe that ERP patterns are found based on the determination of syntactic and semantic nature violations.

Therefore, findings from Harris *et al.* (2000), in sentences with syntactic violation, showed that there was a comparison of the syntactic anaphor agreeing/disagreeing and, in sentences with nonsyntactic violation, there was a comparison of the logophoric anaphor agreeing/disagreeing. The author's prediction was that, in sentences containing a syntactic violation, a P600 would be found, since it has a syntactic nature and causes a similar pattern to that encountered in these types of violations. In logophoric sentences, however, P600 would not be found, since this violation would have a semantic-pragmatic nature. The author also states that it would be difficult to know what type of wave to expect from logophoric violations, given that they are not defined in literature in general. The author expected only to distinguish

¹⁰ Electroencephalograph

¹¹ Event-related brain potentials.

them from violations of syntactic pattern. For this purpose, a study was constructed with 24 sets of experimental phrases, constructed according to the following model:

- (30) a. *The boys' cousin introduced Suzie and himself at the wedding.*
b. *The boys' cousin introduced Suzie and themselves at the wedding.*
c. *The boys' cousin introduced himself at the wedding.*
d. *The boys' cousin introduced themselves at the wedding.*

The task performed by the 40 volunteers consisted of reading the sentences and determining who the recipient of the sentence action was, based on number agreement. Two answer options were shown on the screen, and all subjects were encouraged to anticipate the question and the answer.

The results confirmed that the comparisons involving syntactic constraints exhibited a P600, while constraints with logophors did not. It has been suggested that comparisons involving arguments are mediated by processes sensitive to syntactic constraints, whereas those with logophors are insensitive or sensitive to these constraints in a different manner.

In the attempt to understand sentence processing in real time, focusing on reflexivity, and using as its means the discussion about syntax-only field, on the one hand, and about syntax + discourse on the other, Petra Burkhardt (2002) conducts a study in the English language using the cross-modal paradigm to determine processing in sentences containing logophors. That is, both postures were examined from the perspective of processing.

Because the interpretation of logophoric reflexives involves the access to nonsyntactic operations, such as the access to the notions of discourse and point of view, and the coargument reflexives involve only syntactic operations, the author proposes the hypothesis that the contrast between these two reflexives would have different processing costs.

In the syntax-only position, the two reflexives proposed in the study would have a similar behavior; while in the syntax + discourse position, there would be a contrast between them.

Two tasks were performed, the sentence comprehension task and the lexical decision task, which consisted of presenting the sentence, which the subject heard and, afterwards, answered the comprehension question about it, at random points on the course of the experiment. At a given moment during the sentence presentation, a probe word was shown, and the subject had to decide, by pressing either the button “yes” or the button “no”, whether the probe word was an English word or not. The reaction time taken for the lexical decision was recorded.

For the study, 25 pairs of experimental sentences and 119 distracter sentences were created. In the experimental sentences, each pair consisted of a sentence with a coargument reflexive and a logophoric reflexive, as shown below:

- (31) a. *The woman_i who was arrogant praised^{PROBE} herself_i^{PROBE} because the network had called about negotiations for a leading role.*
 b. *The girl_i sprayed bug repellent around^{PROBE} herself_i^{PROBE} because there were many mosquitoes in the Everglades.*

The sentences with logophoric reflexive included direct objects (*bug repellent*) and prepositions of place (*around* or *behind*). The verbs in both sentences were controlled by frequency, and the total length of the sentences, as well as the distance between the antecedent and reflexive, were paired.

The probe words were placed in two positions: before the reflexive (control position) and after the reflexive (experimental position). They were not semantically related to the sentence, nor did they allow for generating a sentence continuation. In the experimental position, it is possible to measure the resources required while processing the reflexive, determining whether there is an additional cost or not.

By means of the comparison between the times taken for the lexical decision under both conditions, it is possible to determine if there are indications for the syntax-only position (no difference between both conditions) or for the syntax + discourse position (there is a difference between both conditions, with an increase in time for the conditions with logophoric reflexives).

Results have not shown a significant difference for the probe word in the control position, but there was a significant time difference for the probe reaction times (RT) in the experimental position, which recorded a longer time for the logophoric reflexives, thus indicating that the interpretation of logophoric reflexives is costlier to the processor than that of coargument reflexives.

The obtained data are in favor of a syntax + discourse approach to reflexivity, thus suggesting that logophoricity goes beyond syntax. The difference between the coargument reflexives (whose interpretation is merely that of syntactic reflexives) and logophoric reflexives (which implies that their interpretation requires access to syntactic and nonsyntactic information) indicates that there is an additional cost to the processor. This cost, in turn only compatible with the syntax + discourse position that claims that the interpretation of logophoric reflexives requires access to discourse information (extra-syntactic).

These results obtained by Pietra Burkhardt (2002) corroborate Harris et al.'s (2000) ERP findings, as described above, in which the "errors" involving the two types of reflexives produce different event-related brain potential activation patterns, suggesting the existence of syntactic and extra-syntactic processes.

The experiments conducted in English, as described above, are of utmost importance for the study of logophors, since the pragmatic factors have received great attention from the theoretical literature²¹. However, no quantitative studies have been found that determine the acceptability of sentences with logophors. Gibson and Fedorenko (2013) mentioned that the acceptability judgment task is the ideal form for probing syntactic/semantic knowledge, in comparison with other tasks or reading measures.

The experiments in Experimental Psycholinguistics with focus on the processing of logophors are mostly based on Reinhart and Reuland's (1993) perspective, in which the anaphor structural position determines its logophoric nature. In the present work, Reinhart & Reuland's (1993) perspective will also be used as a basis, given that logophoricity will be observed from the anaphor's and the pronoun's structural position.

4 The experiment

Considering the different theoretical approaches to logophoricity in several languages, the experiment described below was intended to verify the acceptability of constructions with logophoric anaphor in BP in order to understand whether or not they are present in the grammar of these language speakers, given that Reinhart and Reuland's (1993) approach shows that these structures in which semantics and discourse operate and have a determining influence in the coreferential resolution of logophoric anaphors are grammatical in English. Furthermore, as presented above, these authors also state that, in the contexts in which there are logophoric anaphors, there may also be a logophoric pronoun.

Galves' (2001, p. 132 *apud* BRITO, 2009, p. 56) study shows that, in some BP dialects, the word "*ele*" (he) may accept an anaphoric reading. Therefore, based on Reinhart and Reuland (1993), this study considers the logophoric property to be opposite to the anaphoric property, since the logophoric property is characterized for having an antecedent outside the binding domain, whether within the sentence or not. Therefore, the pronoun "*ele*" (he) in the experiment is called a logophoric pronoun, because it has an antecedent outside the binding domain in the sentences used. Based on this, this study also seeks to determine the acceptability of the logophoric pronoun in order to establish a counterpoint with the acceptability of the logophoric anaphor.

From these approaches, the study seeks to answer the following questions:

1. Are these constructions as acceptable in Portuguese as they are in English?
2. Are logophoric constructions with anaphors and pronouns equally acceptable?

To achieve this aim, an off-line experiment of acceptability judgment was conducted. According to Leitão (2011):

The determinations obtained from off-line experiments provide information on the interpretation (moment of reflection) of sentences or utterances, that is, they are able to capture the reactions to linguistic stimuli when there

has already been an integration of all linguistic levels (phonological, morphological, lexical, syntactic and semantic). (LEITÃO, 2011, p. 223).

As such, the experiment sought to observe whether or not there is acceptability in logophoric construction in BP, according to the Reinhart and Reuland's (1993) studies. Such acceptability may be confirmed by means of the off-line measurement, since the subjects will issue judgment for the sentences read. Thus, the logophoric anaphor (LA) *ele mesmo* (himself) and the logophoric pronoun (LP) *ele* (him) were used in this construction to determine such an acceptability. The hypothesis predicted for this test was that both structures would be deemed accepted by these language speakers, as provided in Reinhart and Reuland's (1993) Reflexivity Theory.

4.1 Method

4.1.1 Participants

The participants in this research were 77 students majoring in Languages at the Federal University of Paraíba (UFPB), Campus I João Pessoa, all native speakers of Brazilian Portuguese, with ages ranging from 19 to 32.

4.1.2 Material

To conduct this experiment, 12 experimental sentences were created, divided into 2 sets. Each set included six sentences with the logophoric anaphor *ele mesmo* (himself) and six sentences with the logophoric pronoun *ele* (him). In addition to these experimental sentences, 24 distracter sentences were inserted, 12 sentences of which were acceptable, while the other 12 were unacceptable. The purpose of these sentences was to ensure the subject did not realize the nature of the sentences that were the research's object of study. A small instruction was also prepared so that the subject would understand the task. The table below shows examples of the experimental sentences and the instruction in the test.

TABLE 2 – Example of experimental sentences in the pilot test

INSTRUCTION	
Observe the sentences below and answer whether you consider them to be acceptable or unacceptable. There are no right or wrong answers. We only want to check your intuitions about the sentences.	
TYPE OF SENTENCE	SENTENCES
Logophoric anaphor	<i>O irmão de Vítor penteou Lucas e ele mesmo com a escova da sua tia.</i> (Vitor's brother combed Lucas and himself with his aunt's brush.) <i>O genro de Mauro cortou José e ele mesmo com a faca na cozinha.</i> (Mauro's son-in-law cut José and himself with the knife in the kitchen.)
Logophoric pronoun	<i>O irmão de Vítor penteou Lucas e ele com a escova da sua tia.</i> (Vitor's brother combed Lucas and him with his aunt's brush.) <i>O genro de Mauro cortou José e ele com a faca na cozinha.</i> (Mauro's son-in-law cut José and him with the knife in the kitchen.) ¹²

Source: Prepared by the authors.

The experiment's dependent variable was the type of judgment (acceptable vs. unacceptable) made under each condition, and the independent variables were the type of logophor (anaphor and pronoun). Based on this, two experimental conditions were obtained: anaphora with logophoric pronoun and anaphora with logophoric anaphor. The antecedent and the anaphora genders (masculine) were controlled, and all logophors were within a coordination.

¹² The experimental sentences used here adopted the model proposed by Reinhart and Reuland (1993).

4.1.3 Procedure

The experimental acceptability judgment technique (*off-line*) was used, and it consists of checking if a given stimulus is acceptable to these language speakers, since the linguist's intuition is insufficient to determine such a judgment. This technique appears to be ideal to test this phenomenon since this judgment provides a piece of information that enables making inferences about the representation of logophoricity in these speakers' grammar.

The pilot experiment was prepared with Google Docs in a computer, and it was submitted to the subjects through the link to access the test. This same link included the instructions to perform the experiment, which were shown on the top part of the file. The participants read the sentences in front of the computer screen and keyboard. After reading each sentence, the participants would have to choose between acceptable and unacceptable options based on their own judgment. The options were beside each sentence, which had a single marking possibility. The sentences were divided into two sets – one sentence with logophoric anaphor “ele mesmo” (himself) and another sentence with logophoric pronoun “ele” (him). Therefore, the subject's task was to read each sentence and choose one option, that is, the subject would have to mark each sentence as “acceptable” or “unacceptable”. The stimuli presented in the experiment were randomized. As previously mentioned, the intent was to determine which sentence would be acceptable and which one would be unacceptable.

4.2 Results and discussion

The results obtained in this first test did not fully corroborate our original hypothesis formulated. In general, the sentences considered as acceptable were the ones that contained the logophoric anaphor, and the sentences considered as unacceptable were those with the logophoric pronoun. The logophoric anaphor was judged as acceptable 276 times, while the logophoric pronoun was judged as acceptable 213 times. However, no significant result was obtained in the latter, which contradicted the hypotheses that both structures are grammatical in BP.

The results of the judgments made under each condition may be expressed according to the table below:

TABLE 3 – Results of acceptability judgments
for each experimental condition

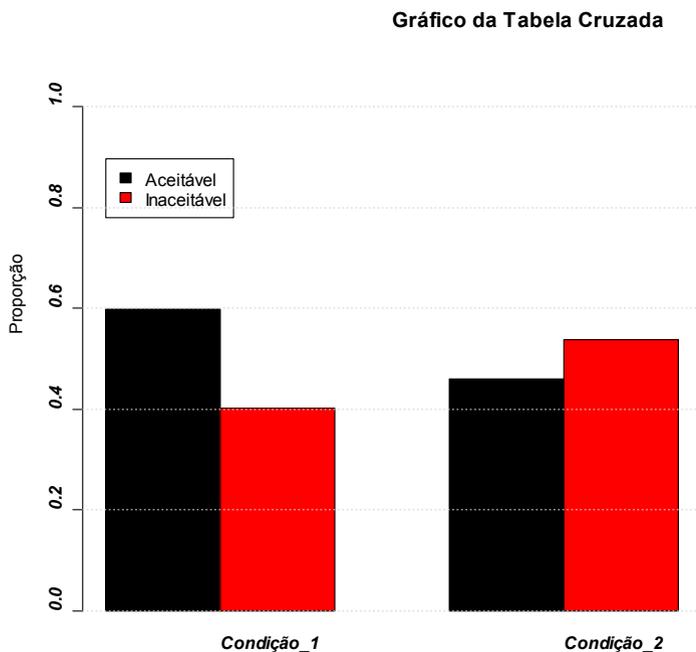
	Condition 1 (anaphora with logophoric anaphor)	Condition 1 (anaphora with logophoric pronoun)	Total
Acceptable	276	213	489
Unacceptable	186	249	435
Total	462	462	924

Source: Prepared by the authors.

Based on the table above, it is possible to observe that the logophoric anaphor obtained a higher number of “acceptable” answers (condition 1), while the logophoric pronoun obtained a higher number of “unacceptable” answers (condition 2).

The difference for the first condition, in which 276 acceptable answers were obtained and 186 unacceptable answers were obtained, was considered statistically significant in the Chi-square proportion test: $\chi^2(1, 462) = 17.532, p < 0.05$. This difference indicates an acceptability of this type of construction in Brazilian Portuguese. For the second condition, with 213 acceptable answers and 249 unacceptable answers, no significant result was obtained using the same statistical test: $\chi^2(1, 462) = 2.805, p = 0.09$. This result does not allow one to identify that logophoric pronouns are unacceptable in this language. Among the experimental conditions used, the Chi-square test of homogeneity provided us with a significant result ($p < 0.05$), indicating that the subjects viewed logophoric anaphor (LA) and logophoric pronoun (LP) as different structures. The chart below expresses this result, showing a reverse trend in the LP condition, when compared to the LA condition:

FIGURE 1 – Chart with the results of the acceptability judgment



Source: Prepared by the authors.

These results are capable of showing us the need to conduct a different experiment in order to explore the study of logophors in BP in greater depth, given that these data previously suggest that the grammar licenses the logophoric anaphors and sees them as different from the logophoric pronouns, which is different from what occurs in the English language (REINHART; REULAND, 1993). It is possible that the reason for this is the test design not having contributed to the acceptance of sentences with logophoric pronoun, since the test may have been biased, given that the subjects who read the sentences with logophoric anaphors also read the sentences with logophoric pronouns. Therefore, to eliminate this possibility, the next step in this research is to conduct an experiment of acceptability judgment whose design is structured so that the conditions are seen by different subjects.

5 General discussion

The present study's hypothesis was to determine whether or not constructions with logophoric anaphors and logophoric pronouns are accepted by BP speakers. Although our intuition may be used to consider such an acceptability, the reality to be checked by means of the data could be opposite, which led us to perform the previously described experiment. As Gibson and Fedorenko (2013) show, the linguist's intuition is not sufficient enough to determine such an acceptability.

The present study points to the reliability of sentence judgment tests and their importance in the study of these structures, given that they do not allow the researcher to assume the acceptability of structures seen as ungrammatical by naive subjects. Hence, this type of test goes beyond the mere use of the researcher's intuition to investigate the BP grammar. After all, if this acceptability study had not been conducted, both structures' acceptability could have been assumed, which was not supported by the acceptability test. The described experiment showed that the subjects considered sentences with logophoric anaphor to be acceptable, and sentences with logophoric pronoun to be unacceptable.

The data obtained based on the acceptability judgment, as conducted in this research, demonstrated that the structures of sentences with logophoric anaphor are considered to be acceptable by native speakers of BP. This type of structure will be considered, as in Reinhart and Reuland (1993), to be grammatical. The data found point toward this possibility, since significant values have been obtained for the acceptability of these logophoric structures.

Under the experimental conditions of logophoric anaphors, the subjects judged them to be acceptable; while in the conditions of logophoric pronouns, no clear significance difference was observed.

These results are capable of showing us a path to be explored in the study of logophors in BP, given that grammar licenses logophoric anaphors and considers them to be different from logophoric pronouns, unlike languages such as English (REINHART; REULAND, 1993).

A possibility of explanation for this difference is the occurrence of bias, since the subjects who read the sentences with logophoric anaphor also read the sentences with logophoric pronoun. Therefore, to exclude this possibility, another acceptability should be conducted with

both conditions: logophoric anaphor and pronoun, following the same model of the test performed, but exposing the subjects to only one of the experimental conditions.

Finally, the Reflexivity Theory predicts that logophoric structures are completely acceptable by English speakers. Therefore, based on the experimental technique used, the conclusion reached is that the data analyzed here show that logophoric anaphors are acceptable. Based on the perception of the subjects participating in the experiment, it has been detected that such a structure most likely composes the grammar of BP speakers.

6 Final considerations

Linguistic studies, especially in the linguistic processing field, which investigate the anaphoric process, seen as relevant to the Binding Theory binding principles in coreference resolution. These studies take these binding constraints' actions into account at the processing moment, questioning the grammaticality of antecedents in coreferential resolution.

The study described here seeks to understand structures with logophoric anaphors that are not addressed by this theory and to compare them with the logophoric pronouns. For this purpose, this study resorts to Reinhart and Reuland's (1993) Reflexivity Theory as a theoretical basis for the conducted experiment. This work seeks to prove the hypothesis that sentences with logophors are acceptable in BP, given that, according to Reinhart and Reuland (1993), these sentences should be considered to be grammatical in languages such as English.

The experiment results illustrated that the logophoric anaphor *ele mesmo* (himself) is acceptable in BP ($p < 0.05$); however, no significant result was obtained for the logophoric pronoun *ele* (he) ($p = 0.09$).

Since this is one of the first studies related to logophoricity in BP, there is still much to be explored. This topic still requires a broader approach in the fields of both Linguistics and Psycholinguistics, given that in this work only the acceptability of two of the different structures with logophors were studied. Therefore, in addition to investigating its acceptability, it is necessary to understand how the processing of these structures occurs, as well as to investigate experimentally the processing of the syntactic anaphor and the logophoric anaphor, with the intention of

determining whether or not these structures differ in terms of processing. It is worth mentioning that a new investigation is warranted to identify how the processing of these semantic-discursive factors takes place in structures with logophors, as well as to observe if the processing of a logophoric element will result in a higher operational cost.

Hence, the representation distinction between the syntactic and logophoric anaphors will be the observed in subsequent steps of this research.

References

- ARIEL, M. *Accessing Noun Phrase antecedents*. London: Routledge, 1990.
- BADECKER, W.; STRAUB, K. The Processing Role of Structural Constraints on the Interpretation of Pronouns and Anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, American Psychological Association, v. 28, n. 4, p. 748-769, 2002. <https://doi.org/10.1037/0278-7393.28.4.748>.
- BRITO, D. B. S. *O se reflexivo no português brasileiro*. 2009. Tese (Doutorado em Linguística) – Faculdade de Letras, UFAL, Maceió, 2009.
- BURKHARDT, P. Logophors: Looking Outside of Syntax; Evidence from Real-Time sentence Comprehension. In: THRIFT, E. *et al.* (Ed.). *Proceedings of Console IX, 2002*. Disponível em: <<http://www.hum2.leidenuniv.nl/pdf/lucl/sole/console9/console9-burkhard.pdf>>. Acesso em: 10 abr. 2015.
- CANTRALL, W. R. *On the nature of the reflexive in English*. 1969. Dissertação (PhD) - Universidade de Illinois, Urbana-Champaign, 1969.
- CANTRALL, W. *Viewpoint, Reflexives, and the Nature of Noun Phrases*. The Hague: Mouton, 1974.
- CHOMSKY, N. *Lectures on government and binding*. Dordrecht: Foris, 1981.
- CHOMSKY, N. *Knowledge of language*. New York: Praeger, 1986.
- CLEMENTS, G. N. The logophoric pronoun in Ewe: its role in discourse. *Journal of West African Linguistics*, West African Linguistics Society, v. 10, p. 141-177, 1975.

FORAKER, S. The processing of logophoric reflexives shows discourse and locality constraints. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, Chicago Linguistic Society, n. 2, p. 295-309, 2003.

GALVES, C. C. *Ensaio sobre gramática do português*. Campinas, SP: Editora UNICAMP, 2001.

GIBSON, E.; FEDORENKO, E. The need for quantitative methods in syntax and semantics research. *Linguagem and Cognitive Processes*, Taylor & Francis Online, v. 28, p. 88-124, 2013.

HARRIS, T.; WEXLER, K.; HOLCOMB, P. J. An ERP investigation of binding and coreference. *Brain and Language*, Elsevier, v. 75, p. 313-346, 2000. <https://doi.org/10.1006/brln.2000.2318>.

KENNISON, S. Comprehending the pronouns her, him, and his: implications for theories of referential processing. *Journal of Memory and Language*, Elsevier, v. 49, p. 335-352, 2003. [https://doi.org/10.1016/S0749-596X\(03\)00071-8](https://doi.org/10.1016/S0749-596X(03)00071-8).

KUNO, S. Functional sentence perspective. A case study from Japanese and English. *Linguistic Inquiry*, MIT Press, v. 3, p. 269-320, 1987.

KUNO, S. Three perspectives in the functional approach to syntax. In: Grossman, Robin E.; JAMES SAN, L.; VANCE, Timothy J. (Ed.). *Papers from the parasession on functionalism*. Chicago: Chicago Linguistic Society, 1975. p. 276-336.

KURODA, S. Where epistemology, grammar and style meet: A case study from Japanese. In: ANDERSON, S.; KIPARSKY, P. (Ed.). *A Festschrift for Morris Halle*. New York: Holt; Rinehart & Winston, 1973.

LEITÃO, M. M. Psicolinguística experimental: focalizando o processamento da linguagem. In: MARTELOTTA, M. E. *Manual de Linguística*. São Paulo: Editora Contexto, 2011.

LEITÃO, M. M.; PEIXOTO, P.; SANTOS, S. Processamento da co-referência intra-sentencial em português brasileiro. *Veredas on-line*, Juiz de Fora, UFJF, v. 2, p. 50-61, 2008.

MENUZZI, S. *Binding Theory and Pronominal Anaphora in Brazilian Portuguese*. 1997. Tese (Doutorado) - Universidade de Leiden, 1997.

NICOL, J.; SWINNEY, D. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, Spring Link, v. 18, n. 1, p. 5-19, 1989.

POLLARD, C.; SAG, I. A. Anaphors in English and scope of Binding Theory. *Linguistic Inquiry*, MIT Press, 23:2, p. 261-303, 1992.

REINHART, T.; REULAND, E. Reflexivity. *Linguistic Inquiry*, MIT Press, v. 24, n. 4, p. 657-720, 1993.

ROSS, J. R. On declarative sentences. In: JACOBS, R. A.; ROSENBAUM, P. S. (Ed.). *Readings in English transformational grammar*. Washington: Georgetown University Press, 1970. p. 222-272.

SELLS, P. Aspects of Logophoricity. *Linguistic Inquiry*, MIT Press, v. 18, n. 3, p. 445-479, 1987.

STURT, P. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, Elsevier, v. 48, p. 542-562, 2003. [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3).

WARSHAWSKY, F. *Reflexivization I & II*. Indiana University Linguistics Club, 1965. [Reproduzido em notas de linguistic underground (*Syntax and Semantics*, v. 7), editado por James D. McCawley. New York: Academic Press, 1976, p. 63-84].

ZRIBI-HERTZ, A. Anaphor binding and narrative point of view: English reflexive pronouns in sentence and discourse. *Language*, Washington, v. 65, n. 4, p. 695-727, 1989. <https://doi.org/10.2307/414931>.

ATTACHMENT 1

The experimental conditions and the sentences used in the acceptability judgment test are presented below:

Anaphora with logophoric anaphor
<p style="text-align: center;"><i>O irmão de André machucou Davi e ele mesmo na BR230.</i> (André's brother hurt Davi and himself on BR230.)</p> <p style="text-align: center;"><i>O genro de Mauro cortou José e ele mesmo com a faca na cozinha.</i> (Mauro's son-in-law cut José and himself with the knife in the kitchen.)</p> <p style="text-align: center;"><i>O primo de Leandro barbeou Arthur e ele mesmo no banheiro da rodoviária.</i> (Leandro's cousin shaved Arthur and himself at the bus station's rest room).</p> <p style="text-align: center;"><i>O irmão de Vitor penteou Lucas e ele mesmo com a escova da sua tia.</i> (Vitor's brother combed Lucas and himself with his aunt's brush.)</p> <p style="text-align: center;"><i>O amigo de Joaquim feriu Tiago e ele mesmo com o brinquedo pontiagudo.</i> (Joaquim's friend hurt Tiago and himself with the sharp toy.)</p> <p style="text-align: center;"><i>O neto de Breno olhou Murilo e ele mesmo no espelho da sala.</i> (Breno's grandson looked at Murilo and himself on the living room mirror.)</p>
Anaphora with logophoric pronoun
<p style="text-align: center;"><i>O irmão de André machucou Davi e ele na BR230.</i> (André's brother hurt Davi and him on BR230.)</p> <p style="text-align: center;"><i>O genro de Mauro cortou José e ele com a faca na cozinha.</i> (Mauro's son-in-law cut José and him with the knife in the kitchen.)</p> <p style="text-align: center;"><i>O primo de Leandro barbeou Arthur e ele no banheiro da rodoviária.</i> (Leandro's cousin shaved Arthur and him at the bus station's rest room).</p> <p style="text-align: center;"><i>O irmão de Vitor penteou Lucas e ele com a escova da sua tia.</i> (Vitor's brother combed Lucas and him with his aunt's brush.)</p> <p style="text-align: center;"><i>O amigo de Joaquim feriu Tiago e ele com o brinquedo pontiagudo.</i> (Joaquim's friend hurt Tiago and him with the sharp toy.)</p> <p style="text-align: center;"><i>O neto de Breno olhou Murilo e ele no espelho da sala.</i> (Breno's grandson looked at Murilo and him on the living room mirror.)</p>

Efeitos de distância linear e marcação no processamento da concordância verbal variável no PB

Linear distance and markedness effects in variable subject-verb agreement processing in BP

Mercedes Marcilese

Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais / Brasil
Núcleo de Estudos em Aquisição da Linguagem e Psicolinguística (NEALP)
mercedes.marcilese@ufjf.edu.br

Erica dos Santos Rodrigues

Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
Laboratório de Psicolinguística e Aquisição da Linguagem (LAPAL)
ericasr@puc-rio.br

Marina Rosa Ana Augusto

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
Laboratório de Psicolinguística e Aquisição da Linguagem (LAPAL)
marinaaug@uerj.br

Késsia da Silva Henrique

Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais / Brasil
Núcleo de Estudos em Aquisição da Linguagem e Psicolinguística (NEALP)
kessiasilvahenrique@gmail.com

Resumo: Este artigo investiga o papel de distância linear entre sujeito e verbo e de marcação morfológica de número no sujeito e no verbo no processamento da concordância verbal variável no português brasileiro

(Ex.: *Policiais Militares, após denúncia, prendeØ/ prendem traficante*). Tomando como ponto de partida estudos sobre o processamento da concordância – baseados em dados de compreensão e de produção dos denominados erros de atração (Ex.: *O tecido das cortinas rasgaram*) – e resultados de pesquisas sobre a concordância variável realizadas no contexto da Sociolinguística Variacionista, foi elaborado um experimento de leitura conduzido por meio de uma *maze task*. Os resultados obtidos sugerem que a distância tem impacto na computação dos traços de número no verbo: tempos de reação significativamente menores foram registrados nas condições de distância longa. Além disso, a marcação de número do sujeito (singular/plural) e do verbo também se mostrou relevante. Sujeitos plurais parecem ser mantidos na memória de forma mais robusta e estabelecer restrições mais pesadas na computação da concordância do que sujeitos singulares, mesmo no contexto de falantes expostos à variação linguística em que marcas não redundantes são atestadas.

Palavras-chave: concordância variável; processamento da concordância; distância linear; marcação morfológica.

Abstract: This paper investigates the role of linear distance between subject and verb and number marking for the processing of Brazilian Portuguese variable subject-verb agreement (e.g. *Policiais Militares, após denúncia, prendeØ/ prendem traficante*). On the basis of production and comprehension studies about attraction errors (E.g. *O tecido das cortinas rasgaram*) and the findings of Sociolinguistics, we conducted a reading experiment making use of a maze task. The results suggest that distance has an impact on the computation of the number features on the verb: with shorter reaction times under longer distance conditions. Results concerning number marking (singular/plural) were also significant: plural subjects representations appear to be more persistent in memory and impose heavier constraints on agreement computation in comparison to singular subjects, even for speakers in the context of linguistic variation concerning non-redundant marking.

Keywords: variable agreement; agreement processing; linear distance; markedness.

Recebido em: 9 de dezembro de 2016.

Aprovado em: 20 de janeiro de 2017.

1 Introdução

Os estudos sobre o processamento da concordância na produção e na compreensão têm se consolidado a partir da investigação dos chamados lapsos de atração. No âmbito dessa literatura, tem sido avaliado um conjunto de fatores que, potencialmente, poderiam interferir na computação das relações de concordância entre o sujeito e o verbo. Sabe-se assim que lapsos ou erros de atração são favorecidos em situações em que o núcleo do sujeito é singular e há uma distância linear e/ou hierárquica entre este e o verbo, como no exemplo (1) retirado de Rodrigues (2006, p.15):

- (1) A análise dos **resultados** experimentais **indicaram** um efeito principal de número do núcleo interveniente no processamento da concordância.

Note-se que os lapsos consistem em falhas de processamento, não estando, pois, associados a um desconhecimento das regras da gramática da língua e, nesse sentido, refletem uma questão de desempenho e não de competência linguística. São, pois, como coloca Fromkin (1973, p.217) “inovações linguísticas não-intencionais”, “desvios do desempenho no que tange à intenção [de fala] fonológica, gramatical ou lexical do falante”.¹

¹ No âmbito da Psicolinguística, lapsos de fala são caracterizados como produções decorrentes de falhas em alguma das etapas do processamento linguístico típico (i.e. em indivíduos sem comprometimento linguístico). Esse tipo de ocorrência foi inicialmente investigado por Meringer e Mayer (1885), dando origem à publicação de uma lista de erros de fala e de escrita no alemão. Na mesma época, Freud em *Psicopatologia da vida cotidiana* analisa lapsos de fala numa perspectiva psicanalítica (FREUD, 1901). Vale salientar que o próprio Freud faz referência – no capítulo V dessa obra – ao mencionado trabalho de Meringer e Mayer. Na década de 1970, a investigação dos lapsos vai ser retomada por Fromkin (1971, 1973) e Garrett (1975, 1976) que propuseram modelos nos quais buscaram estabelecer as rotas do pensamento à fala articulada a partir de dados de erros de fala. No *site* do Max Planck Institute (https://www.mpi.nl/dbmpi/sedb/sperco_form4.pl) é possível acessar uma base de dados de lapsos de fala que inclui *corpora* de várias línguas (inglês, francês, italiano e alemão). Lapsos podem ser categorizados de acordo com as unidades linguísticas envolvidas em cada caso (traços fonológicos, fonemas, sílaba, morfema, palavra, frase ou sentença) e o tipo de mecanismo envolvido no erro (fusão, substituição, adição, apagamento de unidades etc.).

Em línguas em que é registrada variação linguística no que diz respeito à realização da concordância, como é o caso do Português Brasileiro (PB), nem sempre é fácil distinguir entre uma situação típica de ocorrência de lapso e contextos em que a ocorrência de uma variante não padrão seria possível, como ilustrado em (2), em que o núcleo do sujeito é plural e há um potencial elemento atrator entre o sujeito e o verbo.

- (2) **Contas vazias** do Estado **deixa**Ø pagamento de servidores, pela Justiça, sem prazo para acabar. (*Jornal Extra* online, 13/09/2016)

A pergunta que se coloca então é se a ausência de marca de número plural no verbo corresponde a um erro de atração ou, diferentemente, reflete a existência de regras de concordância variável na língua.

Desde os anos de 1970, estudos descritivos conduzidos no âmbito da Sociolinguística Variacionista têm apontado de forma sistemática que a realização da concordância de número no PB constitui um fenômeno variável (LEMLE; NARO, 1977; NICOLAU, 1984; GRACIOSA, 1991; SCHERRE; NARO, 1993, 1997, 1998a, 1998b, dentre outros; BRANDÃO; VIEIRA, 2012; VIEIRA, 2015; MENDES; OUSHIRO, 2015).

Dois regras gerais de realização da concordância de número podem ser identificadas na língua: (i) marcação morfológicamente redundante, com reiteração da informação de plural em todos os itens relevantes envolvidos; (ii) marcação morfológicamente não redundante, que no caso da relação entre sujeito e verbo, seria obrigatoriamente codificada no sujeito, podendo ser omitida no verbo. A regra (ii) é exemplificada em (3) a seguir:

- (3) [...] as criançaØ iráØ aprender os relacionados ao ambiente []²

A variação entre as duas regras levantadas não seria livre, mas condicionada pela atuação de fatores linguísticos e extralinguísticos. No caso específico da concordância entre sujeito e verbo, dentre as variáveis externas que parecem ter algum papel na ocorrência da concordância verbal variável, idade, local de procedência e sexo são alguns dos fatores

² Exemplo retirado de um conjunto de trechos de sentenças aleatoriamente selecionados, coletados pelas autoras de forma anedótica a partir de trabalhos escritos produzidos no contexto de disciplinas diversas, por alunos universitários, entre 2014-2016.

que têm sido investigados. No entanto, o nível de escolaridade é, dentre os fatores sociolinguísticos explorados na literatura, o que tem sido mais frequentemente identificado como relevante (VIEIRA, 1994; VOTRE, 2013).

No que diz respeito às variáveis linguísticas arroladas como relevantes pelos estudos de *corpora* podemos destacar: (i) a posição do sujeito (anteposto/posposto); (ii) os traços semânticos do sujeito (em especial, o contraste entre sujeitos animados/inanimados); e (iii) a distância entre o sujeito e o verbo (COSTA, 1994; SILVA, 2005; SANTOS, 2010; SCHERRE; NARO; 1998b, 1997).

Observa-se assim que um mesmo fator, qual seja, a distância linear entre sujeito e verbo, parece ser relevante tanto para prever a ocorrência de erros de atração quanto para favorecer a ocorrência de concordância não redundante no caso do PB.

É importante salientar que, enquanto as pesquisas sobre processamento da concordância a partir da análise de lapsos é bastante rica, incluindo estudos que investigam a produção escrita (BOCK; MILLER, 1991; BOCK; CUTTING, 1992; BOCK; EBERHARD, 1993; FAYOL *et al.*, 1994; VIGLIOCCO; NICOL, 1998; FRANCK *et al.*, 2002; no PB, RODRIGUES, 2005a, 2005b, 2006, COSTA, 2013; ALMEIDA, 2016; lapsos na escrita, FAYOL *et al.*, 1994; dentre outros),³ trabalhos envolvendo o processamento da concordância variável ainda são escassos e muito recentes (sobre o PB, MARCILESE *et al.*, 2015; HENRIQUE, 2016; AZALIM, 2016; AZALIM *et al.*, submetido; no inglês: SQUIRES, 2014).

³ No caso de pesquisas em processamento conduzidas com falantes de PB, cumpre destacar, no que tange ao diálogo entre Psicolinguística e Estudos de Variação Linguística, os trabalhos de Costa (2014) e Almeida (2016), os quais, embora não tomem a concordância variável como tópico de investigação, trazem algum nível de discussão sobre possível efeito de variação linguística no processamento da concordância. Costa analisa estruturas em que verbos meteorológicos inseridos em orações relativas concordam em número com o antecedente da relativa (Ex. *Essa roupa é perfeita para locais que ventam muito*), e discute as fronteiras entre lapso e estruturas inovadoras no PB. Almeida, por sua vez, ao examinar, na produção e compreensão, concordância em estruturas predicativas, discute efeitos de marcação morfológica no processamento dessas configurações, observando um possível efeito de formas não marcadas (singular e masculino) em estruturas predicativas.

No presente artigo, explora-se – por meio de metodologia experimental – em que medida a distância linear é um fator que se mostra relevante no processamento da concordância variável por falantes universitários.⁴ Assumindo como pressuposto inicial que a concordância não redundante pode não ser a variante mais frequente para esse grupo – em função do nível de escolaridade que amplia as possibilidades de contato com a variante redundante da concordância –, buscou-se verificar se uma maior distância entre sujeito e verbo poderia melhorar a aceitação de sentenças nas quais o sujeito é marcado no plural e o verbo aparece no singular.

Pelo fato de se tratar de um assunto praticamente inexplorado na literatura psicolinguística, serão retomados aqui estudos que avaliaram o papel da distância (linear e hierárquica) e da marcação morfofonológica (singular/plural) no sujeito na ocorrência de lapsos de concordância. Embora os erros de produção não sejam o foco desta pesquisa, essa bibliografia resulta essencial na análise e interpretação dos dados aqui obtidos experimentalmente no que tange ao processamento da concordância verbal variável. Interessa também a discussão mais recente sobre falibilidade do *parser* em relação a fenômenos linguísticos que podem induzir ilusões gramaticais (PHILLIPS; WAGERS; LAU, 2011). Pesquisas conduzidas no âmbito da Sociolinguística Variacionista fornecem ainda subsídios fundamentais para a elaboração da pesquisa da qual este artigo é resultado.

Em § 2 retomamos pesquisas sociolinguísticas que investigaram, em função de análises de *corpora*, a atuação da variável *distância* na alternância das regras de marcação morfofonológica de plural no verbo. Estudos sobre o processamento da concordância que avaliam tanto a produção quanto a compreensão são discutidos em §3. O experimento conduzido é reportado em §4 e as considerações finais são apresentadas em §5.

⁴ O presente artigo é resultado de uma pesquisa conduzida no âmbito do projeto *Interfaces internas e externas na aquisição e no processamento adulto de L1 e L2: concordância e tópico/foco no PB*, financiado pela FAPEMIG (Edital 1/2015 Demanda Universal – Processo APQ-00988-15). O referido projeto conta com a aprovação do Comitê de Ética da Universidade Federal de Juiz de Fora (CAAE: 44123015.6.0000.5147).

2. Concordância verbal variável no PB e distância entre sujeito/verbo

Um conjunto de estudos conduzidos no contexto da Sociolinguística Variacionista tem apontado para a importância do fator distância linear para a realização da concordância verbal no PB (BRANDÃO; VIEIRA, 2009; SCHERRE; NARO, 1997; GRACIOSA, 1991; dentre outros). Segundo Naro (1981), quanto mais nítida for a relação entre sujeito/verbo e/ou quanto mais próximo do verbo estiver o sujeito a que ele se refere, maiores serão as chances de o verbo receber a marcação morfofonológica de plural. Na escrita, inclusive em contextos tidos como mais monitorados – como no caso de textos jornalísticos –, são frequentes as ocorrências de concordância verbal não redundante quando o sujeito está distante do verbo (4-5).

- (4) Os policiais militares, Sargento Souza e Sargento Ottoni (na foto acima), sob o comando do capitão Flávio, **realizou**Ø nessa quarta-feira, 20/01, duas ocorrências que **resultou**Ø em prisões dos autores (dados retirados de HENRIQUE, 2016)
- (5) Policiais Militares, após denúncia, **prende**Ø traficante do bairro Boa União. (dados retirados de HENRIQUE, 2016)

Especificamente no que diz respeito à modalidade escrita, Motta (2011) investigou a ocorrência de concordância verbal variável em redações de alunos frequentando cursos da Educação de Jovens e Adultos (EJA) e alunos matriculados em cursos regulares dos ensinos fundamental e médio em escolas públicas e privadas localizadas na cidade do Rio de Janeiro. O estudo buscou analisar os fatores linguísticos que mais influenciam a marcação de plural de 3ª pessoa nas redações, além de avaliar fatores extralinguísticos como nível de escolaridade (EJA, ensino fundamental e médio), sexo dos participantes e faixa etária. Os dados coletados foram divididos em três grupos: o grupo 1 formado por redações de alunos de uma escola da rede privada, o grupo 2 por produções de alunos de uma escola da rede Municipal de ensino da cidade do Rio de Janeiro, e o grupo 3 por produções de alunos de uma escola da rede Estadual de Ensino do Estado do Rio de Janeiro. A análise

conduzida sugere que, na modalidade escrita, a saliência fônica⁵ seria um fator linguístico decisivo para presença ou ausência da marca de plural em verbos da terceira pessoa. Já a relevância da distância do verbo em relação ao sujeito foi verificada em algumas redações; entretanto, foi registrado um número reduzido de ocorrências (19 dados, com 12 ocorrências com marcação de plural, totalizando 63% do total). O fato de o sujeito estar anteposto ao verbo também se mostrou um fator determinante para a realização da concordância redundante. Segundo a autora, a posição do sujeito anteposto em relação ao verbo favorece consideravelmente o uso da variante padrão na concordância verbal. De modo geral, os dados analisados por Motta (2011) não se mostram conclusivos com relação à relevância da distância linear em virtude do pequeno volume de ocorrências de sujeitos distantes de verbos na amostra analisada. No entanto, vários estudos prévios (BRANDÃO; VIEIRA, 2009; SCHERRE; NARO, 1997; GRACIOSA, 1991) reportam que a distância linear seria um dos fatores determinantes para explicar a ocorrência variável da marcação morfológica de plural nos verbos.

De acordo com Scherre e Naro (1997), quanto mais material interveniente houver entre o núcleo do sujeito e o verbo, menor é a probabilidade de haver concordância redundante entre ambos os elementos. Uma das explicações aventadas para esse fenômeno seriam restrições na capacidade de memória de trabalho dos indivíduos, que teriam maiores dificuldades para recuperar a informação de número do núcleo do DP sujeito.

Graciosa (1991), ao analisar a alternância na marcação morfológica de plural no verbo na fala de falantes cultos cariocas, verifica que: “quando há proximidade linear entre SN e SV há maior garantia da regra se aplicar” (GRACIOSA, 1991, p.69). Santos (2010, p. 98), por sua vez, observa que a concordância verbal não redundante seria condicionada pelo fator “presença de elementos”, enquanto a realização

⁵ A saliência fônica está associada ao fato de formas singulares e plurais terem maior ou menor identidade fônica quando comparadas entre si. Nesse sentido, as formas singular e plural de verbos como *ser*, na terceira pessoa do presente do indicativo (*é/são*), teriam um maior grau de saliência fônica do que verbos do tipo de *comer* (*come/comem* no presente do indicativo), em que as formas singular e plural são mais semelhantes fonética e morfológicamente entre si. Análises de *corpora* sugerem que quanto menor o grau de saliência fônica de um item, mais favorável seria o contexto à não realização redundante da marca de plural (LEMLE; NARO, 1977, dentre outros).

da concordância verbal é condicionada pelo fator “ausência de elementos” entre o sujeito e o verbo. Vale destacar que, enquanto a pesquisa de Graciosa (1991) investigou o fenômeno da concordância verbal variável na produção oral de falantes adultos, a pesquisa de Santos (2010) analisou o cancelamento da marca de plural no verbo em narrativas de crianças frequentadoras de entidades filantrópicas em Maceió.

Em um estudo comparativo recente realizado por Rubio (2012), houve 67,9% de cancelamento da marcação redundante em dados de escrita em sentenças nas quais a distância entre sujeito e verbo era de entre seis e dez sílabas. Segundo o autor, esses resultados: *“demonstram que o alargamento da distância do verbo em relação a seu sujeito promove, conforme premissas pré-apresentadas, enfraquecimento da concordância verbal”* (p.280).

O papel da distância linear na realização da concordância variável tem sido atestado também em outras variedades do português. Lopes e Baxter (2010), considerando neste caso o português de Tonga, apontam para a interferência do fator distância na realização da concordância. Segundo os autores, a complexidade do SN (sintagma nominal) e a distância entre sujeito e verbo condicionariam a realização das regras de concordância nessa variedade do português. De acordo com Lopes e Baxter (2010), em termos cognitivos e pragmáticos, a adjacência entre esses elementos ajuda na configuração e desenvolvimento da regra. Já Rubio (2015), analisando o português europeu, reporta que a posição do verbo com relação ao sujeito (anteposto/posposto) influenciaria fortemente a ocorrência da concordância variável nessa variedade. Em particular, sujeitos antepostos combinados com uma maior distância (mais de 10 sílabas) configuram um contexto que desfavorece a marcação redundante.

3. Estudos sobre o processamento da concordância: os erros de atração e concordância variável

Como foi previamente mencionado, na literatura psicolinguística, o mecanismo da concordância tem sido explorado principalmente no que diz respeito ao que vem sendo chamado de “erros de atração” (CLAHSEN; HANSEN, 1993; VIGLIOCCO; BUTTERWORTH; GARRET, 1996; VIGLIOCCO; NICOL, 1998). Os erros de atração se caracterizam pelo fato de o verbo concordar não com o núcleo do

sujeito, mas com outro núcleo nominal interveniente, em estruturas do tipo ilustrado em (6).

(6) O álbum das **fotos rasgaram**. (RODRIGUES, 2005a, p. 146)

Diversas pesquisas têm reportado uma assimetria entre singular e plural na indução de erros de atração (BOCK; EBERHARD, 1993; EBERHARD, 1999; no PB, RODRIGUES, 2005a). Assim, a condição em que o N1 é singular e o N2 é plural (7) registra um número maior de erros de atração do que a condição em que o N1 é plural e o N2 é singular (8). É importante lembrar que a configuração em (8), mas não em (7), é a que está envolvida nos casos de concordância variável (i.e. sujeito plural + verbo no singular).

(7) A caixa_{N1} das canetas_{N2}... sujou/sujaram.

(8) As contas_{N1} do estado_{N2}... fechou/fecharam.

O efeito da distância linear no processamento da concordância foi inicialmente investigado em estudos de produção por Bock e Miller (1991), em experimento no qual manipularam, entre outros fatores, o tamanho dos preâmbulos que funcionavam como sujeitos das sentenças, a partir da colocação de uma média de dois adjetivos pré-nominais no modificador (Ex.: *The key to the cabinets* vs. *The key to the ornate Victorian cabinets* / *A chave dos armários* vs. *A chave dos [ornamentados vitorianos] armários*). Nesse mesmo experimento, os autores também variaram o tipo de modificador do núcleo do sujeito – sintagma preposicionado (PP) e oração relativa –, e o número tanto do núcleo do sujeito quanto do núcleo nominal do modificador. Não foi observado efeito principal de tamanho do modificador nem são reportados efeitos de interação entre essas variáveis.

Bock e Cutting (1992), em experimento posterior, também envolvendo modificadores preposicionais e oracionais, obtiveram um efeito do número do nome local, do tipo de modificador e também do tamanho do modificador. Houve também efeito de interação entre número do nome local e tipo de modificador e entre tipo e tamanho do modificador. No que tange à distância, foi observado efeito de tamanho do modificador apenas para PPs. Esses resultados foram tomados como evidências na direção da hipótese do “empacotamento oracional”,

segundo a qual informação de número de um núcleo interveniente teria menos chance de induzir erros de atração quando estivesse em uma oração, porque a oração funcionaria como uma espécie de “ilha” que isolaria o núcleo interveniente do núcleo do sujeito, dificultando qualquer tipo de interferência no estabelecimento da concordância.

No que tange à literatura em compreensão, resultados de um estudo envolvendo leitura de estímulos, com emprego da técnica de monitoramento do olhar (*eyetracking*), conduzido com falantes de hebraico (DEUTSCH, 1998) sugerem que a distância linear entre o núcleo do sujeito e o verbo é um fator que afeta a detecção de erros de concordância de número e de gênero no verbo (dado que o verbo concorda em número e em gênero com o sujeito em hebraico). Nessa pesquisa, foram mais facilmente identificados erros na condição distância curta do que longa.

Pearlmutter (2000) conduziu dois experimentos de leitura automonitorada⁶ com vistas a avaliar efeitos de distância no processamento da concordância na compreensão. O autor manipulou a distância entre o elemento interveniente e o núcleo do sujeito, de modo a poder contrastar duas hipóteses acerca de como a concordância seria computada e possíveis efeitos de interferência: (i) uma hipótese baseada na ideia de percolação (VIGLIOCCO; NICOL, 1997) e em um sistema de transmissão hierárquica de traços (*hierarchical feature-passing system*), e (ii) uma hipótese ancorada na proposta de um sistema linear de *slot* de memória (*linear slot-based system*).

De acordo com a hipótese (i), a distância hierárquica entre o elemento interveniente e a projeção máxima do núcleo do sujeito seria o fator determinante de efeitos de interferência. Logo, quanto mais próximo da projeção mais alta estivesse o elemento com número incongruente, maior a chance de este vir a afetar o processamento da concordância.

⁶ O paradigma da leitura automonitorada foi introduzido independentemente por vários pesquisadores diferentes (cf. MITCHELL, 2004) e se caracteriza pelo fato de permitir que o próprio participante controle – pressionando um botão – o tempo de exposição de cada palavra, segmento ou frase inteira durante sua leitura. Assume-se que o tempo levado para pressionar o botão (passando de um segmento, palavra ou frase para o/a seguinte) depende das propriedades daquilo que está sendo lido e está relacionado com o curso dos processos cognitivos durante a leitura e a compreensão.

Assim, no contraste entre as sentenças (9) e (10), abaixo, a primeira geraria mais efeitos de interferência do que a segunda:

- (9) The lamp near the paintings of the house was damaged in the flood.

A lâmpada perto dos quadros da casa foi danificada durante a enchente.

- (10) The lamp near the painting of the houses was damaged in the flood.

A lâmpada perto do quadro das casas foi danificada durante a enchente.

De acordo com a hipótese (ii), por outro lado, efeitos de interferência estariam diretamente associados a esvaecimento de informação na memória de trabalho. A ideia é que haveria uma localização na memória reservada para rastrear o número do NP (*Nominal Phrase* – sintagma nominal) sujeito durante o processamento e que a informação de número do núcleo do sujeito tenderia a decair com o tempo ou distância. Assim, um elemento interveniente mais distante do núcleo do sujeito teria mais chance de gerar interferência, pois a representação de número do núcleo já teria sofrido processo de esvaecimento. Logo, uma sentença como (10) teria mais chance de provocar efeitos de interferência do que (9).

Pearlmutter (2000) observou efeito de interferência apenas no experimento 2, cujos estímulos eram todas sentenças com N1 plural (Ex. *The lamps near the painting(s) of the house(s) were... / As lâmpadas próximas do(s) quadro(s) da(s) casa(s) foram*). Nesse experimento, quando o número do N2 era incongruente em relação ao número do N1, os tempos de leitura no verbo foram maiores.

Os resultados obtidos nesse segundo experimento estão em acordo com a hipótese de que distância hierárquica é fator relevante no processamento da concordância e também com a ideia de que marcação de número do núcleo do sujeito afeta esse processo.

Esses resultados de compreensão vão na mesma direção do observado em estudos de produção (VIGLIOCCO; NICOL, 1998, com falantes de inglês e RODRIGUES, 2006, com falantes de português brasileiro).

Pearlmutter observa, contudo, que não se pode descartar totalmente um possível efeito de distância métrica. Segundo o autor, caso não houvesse alguma interferência desse fator, seria esperado que os sujeitos plurais do experimento 2 fossem tão invulneráveis à interferência como sujeitos plurais em construções envolvendo um único PP (Ex. *The keys to the cabinet / As chaves do armário*). Um ponto adicional levantado por Pearlmutter diz respeito à natureza dos elementos intervenientes entre o núcleo do sujeito e o verbo. Segundo o autor, talvez o que deva ser considerado não é o número de elementos intervenientes puramente, mas sim a possibilidade de estes precisarem ser computados em um processo de rastreamento de traços de número.

Wagers, Lau e Phillips (2009) conduziram um conjunto de 7 experimentos envolvendo estruturas com elementos intervenientes entre o núcleo do sujeito e o verbo, as quais potencialmente poderiam gerar distúrbios no processamento da concordância. Entre as estruturas investigadas, os autores trabalharam com NPs modificados por PPs e com NPs modificados por orações relativas. No caso dos modificadores do tipo PP, o N1 foi sempre usado no singular e foram manipulados o número do N2 e o número do verbo (Ex. *The key to the cell(s) (unsurprisingly) was(were) rusty from many years of disuse / A chave da(s) cela(s) (sem surpresa) ficou/ficaram enferrujadas pelos muitos anos de falta de uso*); no caso dos modificadores oracionais, foram manipulados tanto o número do N1 quanto do N2 e o número do verbo da relativa (Ex. *The musician(s) who the reviewer(s) praise(s) / O(s) músico(s) que o(s) crítico(s) elogia/elogiam...*).

Em relação aos resultados obtidos nesse conjunto de experimentos, é relevante aqui destacar o fato de os autores terem reportado efeitos de atração em configurações de relativas (em que o verbo parece concordar com o sujeito da oração principal – *The musician who the reviewers praises / O músico que os críticos elogiou*) e de só terem sido observados efeitos de atração em sentenças agramaticais. Tais resultados são apontados como evidências para um mecanismo de recuperação dos traços do sujeito orientado por pistas (similar ao modelo de *slot* de memória de Pearlmutter, 2000). Os autores argumentam que, se o problema estivesse associado à representação de número do sujeito (como no modelo de transmissão hierárquica de traços), seria esperado um estranhamento também no caso de sentenças gramaticais (Ex.: *The key to the cells (unsurprisingly) was rusty from many years of disuse / A chave das celas (sem surpresa)*

ficou enferrujada pelos muitos anos de falta de uso), uma vez que o modificador ao afetar o número da projeção hierárquica mais alta seria incongruente com o verbo. Assim, no exemplo dado, o traço de número de *cells* (*celas*), por um processo de percolação, poderia acabar por especificar o número do DP como plural, o que seria incompatível com o número do verbo, singular – *was* (*foi/ficou*). Rodrigues (2011) nota que os autores reportam um efeito de atração, capturado na posição do verbo, em uma das comparações entre pares gramaticais: *The key to the cells unsurprisingly was rusty from many years of disuse* (com N2 plural) vs. *The key to the cell unsurprisingly was rusty from many years of disuse* (com N2 singular/condição controle), com maior tempo de leitura para a primeira. Esse resultado, não esperado segundo a hipótese dos autores, é explicado em termos de uma continuação do efeito de marcação de número que se estenderia para a região seguinte à do advérbio.

Rodrigues (2011), em um experimento *off-line*, que consistiu na adaptação da técnica de julgamento de gramaticalidade,⁷ verificou que a presença de um N2 plural na posição de modificador do núcleo do sujeito reduz expressivamente a rejeição de sentenças em que a concordância entre o sujeito e o verbo é agramatical – enquanto a média de aceitação para a condição N2 singular + V plural (Ex. *O pediatra da criança da creche receitaram o remédio*) foi de 0,5 em um máximo *score* de 4, a média para a condição N2 plural + V plural (Ex. *O pediatra das crianças da creche receitaram o remédio*) foi de 2,3. Já no caso das condições gramaticais (N2 singular + V singular: *O pediatra da criança da creche receitou o remédio* vs. N2 plural + V singular: *O pediatra das crianças da creche receitou o remédio*) não houve diferença em termos da taxa de aceitação, tendo ambas apresentado resultados próximos ao *score* máximo (respectivamente, 3,8 e 3,7).

Os resultados de Rodrigues (2011) são compatíveis com a proposta de Wagers, Lau e Phillips (2009), segundo a qual, no caso da

⁷ O procedimento experimental adotado consistiu na apresentação de sentenças, palavra por palavra, no centro de uma tela de projeção. Os participantes receberam a informação de que as sentenças haviam sido produzidas por falantes estrangeiros e que a tarefa a ser realizada consistia em dizer se o estrangeiro dominava bem o português. Os julgamentos, registrados em um bloco de respostas, poderiam ser de dois tipos: “D” para Domina ou “N” para não domina. Havia um tempo fixo entre a apresentação dos estímulos, e a resposta deveria ser dada da forma mais rápida possível.

compreensão, o efeito de atração ocorreria no recesso à informação acerca do número do sujeito a partir da identificação do verbo. Uma alternativa considerada pelos autores, no âmbito dessa hipótese, é a de que o número do verbo poderia ser antecipado a partir de um mecanismo preditivo. Apenas nos casos de agramaticalidade ocorreria um processo de reanálise. Essa explicação é, em certa medida, compatível com o modelo de Produção Monitorada por *Parser* (PMP), apresentado por Rodrigues (2006), para explicar erros de concordância por atração na produção.

Uma questão que os experimentos sobre processamento da concordância na compreensão levantam é a da infalibilidade do *parser* às chamadas ilusões gramaticais (PHILLIPS; WAGERS; LAU, 2011). Segundo os autores, as *ilusões gramaticais*, de um modo geral, seriam geradas devido a equívocos do *parser* durante o processamento de sentenças em contextos sintáticos bastante específicos, tais como: anáfora, concordância, caso e dependências. De acordo com os autores, apesar de o *parser* humano ser muito eficiente na implementação de determinadas restrições gramaticais complexas, como as restrições no uso de pronomes reflexivos, este apresentaria falhas em restrições relativamente simples, gerando, assim, as ilusões gramaticais.

Como foi mencionado anteriormente, pesquisas relativas ao processamento da concordância variável são bastante escassas na literatura. O trabalho de Squires (2014) constitui um estudo exploratório que busca analisar o papel de um conjunto de variáveis sociais (classe social, sexo e etnia) no processamento da variação relativa à concordância no inglês americano. Especificamente, foi investigado o modo como a variação na concordância entre sujeito e verbo nessa língua é processada durante a compreensão de sentenças. Para isso, foram conduzidos três experimentos de leitura automonitorada. Três padrões distintos de realização da concordância foram investigados, sendo eles:

- a) Concordância padrão (they don't/ he doesn't);
Plural: *After eating, the turtles don't walk very fast.*
[Depois de comer, as tartarugas não caminham muito rápido].
Singular: *After eating, the turtle doesn't walk very fast.*
[Depois de comer, a tartaruga não caminha muito rápido].

- b) Concordância não padrão (he/she don't);
After eating, the turtle don't walk very fast.
- c) Concordância denominada “incomum” pela autora, i.e., uma opção não reconhecida como variação presente no inglês americano (they doesn't).
After eating, the turtles doesn't walk very fast.

Participaram 112 voluntários, todos eles estudantes de cursos de graduação nos Estados Unidos e com conhecimento da norma padrão de concordância. Os participantes responderam um questionário e foram agrupados de acordo com três categorias: classe (alta ou baixa), etnia (branco, afroamericano ou outros) e sexo (masculino e feminino). Tomando como ponto de partida estudos sociolinguísticos prévios, as previsões levantadas pela autora foram que participantes de classe média baixa, afroamericanos e homens seriam menos afetados pelas diferenças nos padrões de concordância, registrando tempos de reação equivalentes nas condições padrão e não padrão. Esperava-se ainda que participantes de classe social mais alta, mulheres e brancos fossem mais sensíveis a essas diferenças, apresentando tempos de reação mais altos para as condições não padrão e incomum.

Em virtude do foco específico da nossa pesquisa, retomaremos aqui apenas os resultados relativos à *performance* dos participantes de forma global (sem considerar os fatores grupais investigados). Independente do grupo social, sentenças com concordância padrão foram lidas mais rapidamente e sentenças com concordância incomum apresentaram os maiores tempos de leitura, enquanto as sentenças com concordância não padrão registraram tempos de leitura intermediários. Em conjunto, os resultados obtidos por Squires (2014) sugerem um processamento diferenciado da forma não padrão de marcação do plural na relação sujeito-verbo quando comparada com a forma padrão.

No que tange ao PB, os trabalhos de Marcilese *et al.* (2015), Henrique (2016), Azalim 2016, Azalim *et al.* (submetido) e Marcilese *et al.* (em prep.) buscaram investigar o processamento da concordância variável nominal e verbal. Marcilese *et al.* (2015) conduziram um experimento de escuta automonitorada com alunos universitários cujos resultados foram na mesma direção do reportado por Squires (2014) para o inglês e revelaram tempos de escuta significativamente maiores nas

condições de concordância não redundante quando comparadas com as condições de concordância redundante (*Os esquilo desceu da árvore* vs. *Os esquilos desceram da árvore*). Resultados semelhantes foram obtidos por Henrique (2016) a partir de um experimento de produção eliciada por repetição que registrou tempos de reação maiores nas condições não redundantes.

Azalim (2016), Azalim *et al.* (submetido) e Marcilese *et al.* (em prep.) investigaram o papel da saliência fônica no processamento da concordância nominal variável no PB. Para tal, foram conduzidos experimentos de produção eliciada por repetição com adultos universitários, alunos da EJA (Educação de Jovens e Adultos) e crianças na faixa dos 6 anos de idade. Novamente, tomados em conjunto os resultados indicam um processamento diferenciado das formas padrão e não padrão, com tempos de reação significativamente maiores nas condições de concordância não redundante e menores médias de respostas-alvo, para todos os grupos investigados (i.e independente de escolaridade e idade).

Diante do quadro traçado por esses estudos prévios, o presente artigo visa a dar continuidade a essa linha de pesquisa que busca avaliar o processamento da concordância variável no PB. No caso do experimento reportado na próxima seção, as questões que se colocam são:

- (i) Considerando-se que os participantes são falantes universitários (e, portanto, com maior exposição – pelo menos na escrita – à concordância redundante), a presença de uma marca de número plural apenas no sujeito (gramatical na variedade não padrão do PB) geraria menor estranhamento do que uma condição agramatical em qualquer variedade do PB (sujeito singular + verbo plural: **A menina dançaram*)?
- (ii) A distância linear, que pode favorecer a ocorrência de erros (na produção) e sua não percepção como lapsos (na compreensão), interferiria no processo de verificação de traços do verbo pelo *parser*, fazendo com que este fosse mais vulnerável a um efeito de ilusão gramatical?
- (iii) A distância pode contribuir para uma maior “aceitação” de uma variante que – não necessariamente – faz parte da gramática dos nossos participantes?

4. Distância linear no processamento da concordância verbal variável

Com o objetivo de investigar o papel da distância linear na ocorrência da concordância verbal variável, foi conduzido um experimento de leitura por meio de uma tarefa de labirinto (*maze task*).⁸ Em linhas gerais, em experimentos que empregam essa técnica, a tarefa do participante consiste em fazer escolhas sobre qual de duas opções é mais adequada para dar sequência a uma sentença. Diferentemente dos experimentos de leitura automonitorada tradicionais, a tarefa de labirinto requer a integração incremental local de cada nova palavra/segmento com o contexto precedente, o que pode contribuir para inibir eventuais efeitos de *spillover*. É importante lembrar que para o participante o foco da tarefa a ser realizada é a construção de uma frase completa com sentido. Dessa forma, consideramos que aspectos pontuais como a marcação morfofonológica dos itens que conformam a frase em questão podem ficar menos salientes do que em tarefas de leitura convencionais. Por esse motivo, embora o fenômeno aqui investigado seja muito frequente na produção oral, mas proscrito na norma padrão escrita, consideramos que uma tarefa de leitura nos moldes da *maze* poderia ser uma opção metodológica que atenderia aos nossos objetivos.

Outro ponto relevante a ser considerado é o fato da tarefa de *maze* ser concebida de modo que apenas uma das opções apresentadas em cada etapa da leitura possibilita a formação de uma sentença completa. Assim, no nosso caso específico, em que o segmento crítico era o verbo da frase (que podia concordar ou não de forma redundante com o sujeito), este sempre constituía a melhor escolha para a construção global da sentença. Pode-se dizer então que nesta tarefa temos uma situação de “escolha forçada” e o que é de fato avaliado é o quanto tal escolha se mostra demandante (em termos de tempo de reação) para o participante. De acordo com esse raciocínio, escolhas mais rápidas indicariam uma integração mais automática e inconsciente das informações, enquanto escolhas mais demoradas revelariam a percepção de algum tipo de incongruência entre o item relevante para dar prosseguimento à frase e o material previamente processado.

⁸ Maiores detalhes sobre essa técnica podem ser encontrados em Forster, Guerra e Elliot (2009) e no *site* <<http://www.u.arizona.edu/~kforster/MAZE/>>.

4.1 Método

As variáveis independentes foram: *distância linear* entre sujeito e verbo (longa, curta e zero), *número do verbo* (singular/plural) e *número do sujeito* (singular/plural). As duas primeiras variáveis foram fatores *within-subjects* e a terceira, fator grupal (*between-subjects*). As variáveis dependentes foram: o tempo de reação na leitura/escolha do alvo (o segmento crítico considerado foi o verbo da sentença) e o número de respostas-alvo (escolha do verbo da sentença).

No presente artigo, reportamos apenas os resultados detalhados relativos à primeira variável dependente, i.e. tempo de reação. O número de respostas-alvo não se mostrou uma medida informativa já que – como previsto pelo *rationale* da técnica anteriormente apresentado – o percentual de escolhas do verbo-alvo foi bastante alto em todos os grupos e condições (Escolhas do alvo – sujeito plural: distância longa 90%, distância curta 83%, distância zero 80%; sujeito singular: distância longa 92%, distância curta 89%, distância zero 87%). Embora possa ser observada uma diminuição do número de respostas-alvo em função da distância (mais respostas para as condições de distância longa do que para as restantes), essas diferenças não foram estatisticamente significativas.

A apresentação dos estímulos experimentais foi feita de acordo com a distribuição em quadrado latino (todos os membros do grupo foram expostos a todas as condições, mas a mesma sentença não foi apresentada em mais de uma condição para o mesmo participante). Ao longo do teste, foram apresentadas 18 sentenças experimentais (3 por condição) e 18 distratoras. Utilizamos um número menor de distratoras do que o habitual (proporção 1 experimental para 2 distratoras) pelo fato de as próprias sentenças experimentais apresentarem uma variação considerável entre si no que diz respeito ao tamanho. Nesse sentido, estimamos que as próprias sentenças experimentais em cada condição funcionariam como distratoras entre si. A tabela 1 apresenta exemplos das condições experimentais. Além disso, a própria natureza da tarefa, que demanda que o participante lide com “material irrelevante extra” a cada escolha, também traz um elemento de distração que não está presente em tarefas de leitura tradicionais.

TABELA 1
Exemplos das condições experimentais

Grupo 1 – Sujeito plural	Condição
Os alunos no início da aula atentamente escutaram a professora.	Suj pl. V pl. longa
Os alunos atentamente escutaram a professora.	Suj pl. V pl. curta
Os alunos escutaram a professora.	Suj pl. V pl. zero
Os alunos no início da aula atentamente escutou a professora.	Suj pl. V sg. longa
Os alunos atentamente escutou a professora.	Suj pl. V sg. curta
Os alunos escutou a professora.	Suj pl. V sg. zero
Grupo 2 – Sujeito singular	Condição
O aluno no início da aula atentamente escutaram a professora.	Suj sg. V pl. longa
O aluno atentamente escutaram a professora.	Suj sg. V pl. curta
O aluno escutaram a professora.	Suj sg. V pl. zero
O aluno no início da aula atentamente escutou a professora.	Suj sg. V sg. longa
O aluno atentamente escutou a professora.	Suj sg. V sg. curta
O aluno escutou a professora.	Suj sg. V sg. zero

Participantes

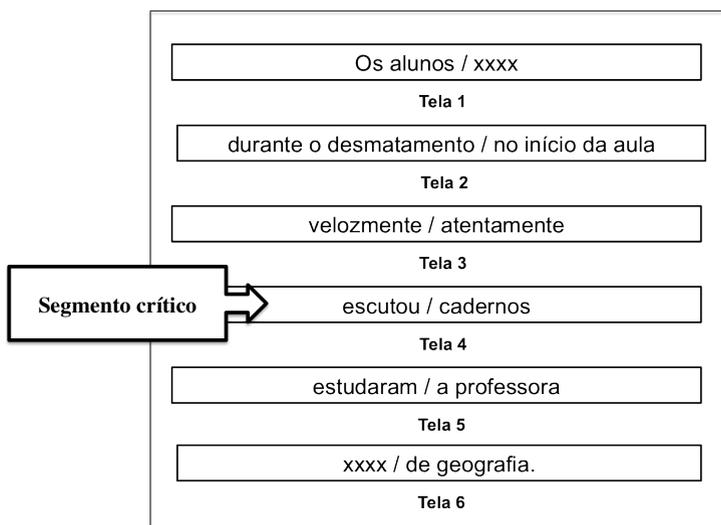
Participaram do experimento 40 adultos voluntários. Todos os participantes eram estudantes universitários de graduação ou pós-graduação, de cursos diversos. A idade média dos participantes foi de 26 anos. Os participantes foram divididos em dois grupos de acordo com a variável grupal *número no sujeito*. Assim, para um grupo só foram apresentadas frases experimentais com sujeitos plurais e, para o segundo grupo, sentenças contendo sujeitos singulares.

Procedimento

Durante a atividade experimental, frases divididas em segmentos são apresentadas na tela do computador. Para cada etapa da tarefa de leitura, duas palavras/segmentos separados por barras são exibidos. Apenas uma das opções apresentadas dá sequência à sentença de forma coerente. Em cada etapa, o participante seleciona uma das duas palavras/

segmentos apresentados utilizando um dos dois botões disponíveis no teclado e marcados respectivamente por ← (esquerda) e → (direita). O primeiro e o último segmento não envolvem escolhas reais (a alternativa nesses casos é indicada por uma sequência de letras x, como ilustrado na Figura 1), já que têm como função iniciar e encerrar cada frase. Após o início da atividade experimental, o participante controla a velocidade de aparição de novos estímulos na tela conforme vão sendo feitas as escolhas entre as duas opções exibidas até completar uma frase (cujo final era graficamente indicado pela presença de um ponto). O experimento foi programado por meio do *software Linger*,⁹ versão 2.94, que também randomizou os itens e captou o tempo de reação e as escolhas dos participantes para posterior análise. Cabe salientar que, no caso do experimento conduzido, no par de escolhas correspondente ao segmento crítico (i.e. aquele que contém o verbo da frase), a alternativa ao verbo nunca era outro verbo (cf. o exemplo apresentado na Figura 1). A duração de cada sessão experimental foi de aproximadamente 12 minutos.

Figura 1 – Exemplo esquemático do procedimento experimental



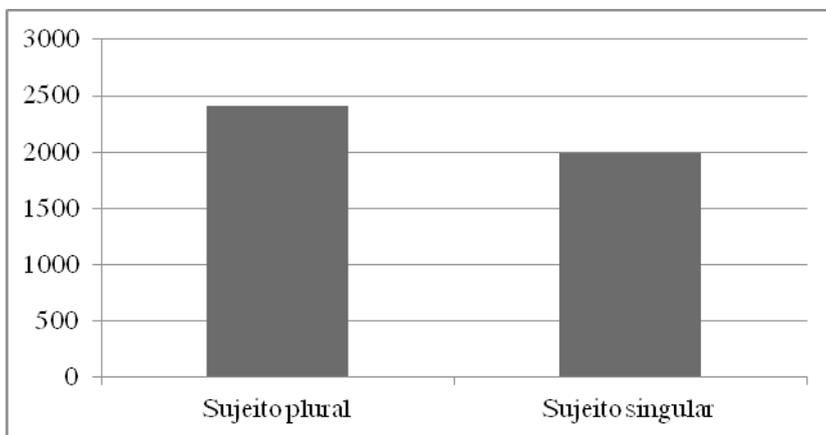
⁹ O *Linger* é uma plataforma flexível para experimentos que investiguem o processamento da linguagem. O *software* foi desenvolvido no laboratório de Ted Gibson e é disponibilizado gratuitamente, funcionando em computadores com qualquer sistema operacional (<<http://tedlab.mit.edu/~dr/Linger/>>).

4.2 Resultados e discussão

Os resultados relativos à variável tempo de reação foram submetidos a uma análise da variância com *design* fatorial 2x3x2 (*número do verbo x distância x número do sujeito*), sendo que a variável *número do sujeito* foi tomada como fator grupal. Para tal, foi utilizado o *software* estatístico gratuito ezANOVA.¹⁰ Cabe destacar que o tempo de reação foi medido no segmento delimitado como crítico, qual seja, o verbo da sentença (cf. Figura 1); todos os resultados reportados a seguir remetem a essa variável dependente.

A análise dos tempos de reação no segmento crítico (i.e. par de opções que contém o verbo) revelou um efeito principal de *número do sujeito*, com médias de tempo de reação significativamente maiores para as sentenças com sujeitos plurais ($F(1,38) = 6.28, p=0.02$) (Médias: 2412.8ms sujeito plural e 1988.5ms sujeito singular).

Gráfico 1 – Tempo médio de resposta (em ms) em função do número do sujeito.

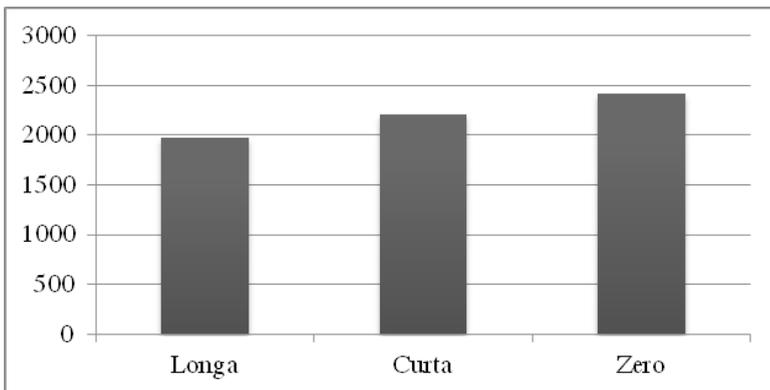


Foi obtido também um efeito principal de *distância*, em decorrência do aumento gradual no tempo de reação em função da distância, com tempos maiores nas condições com distância zero e menores nas condições com distância longa ($F(2,76) = 8.84, p=0.0003$) (Médias: 1979.1ms, 2207.4ms e 2415.6ms para a distância longa, curta

¹⁰ Recurso disponível em: <<http://www.cabiati.com/micro/ezanova/>>. Esse mesmo *software* fornece ainda as comparações entre pares de condições por meio do teste de *Tukey*.

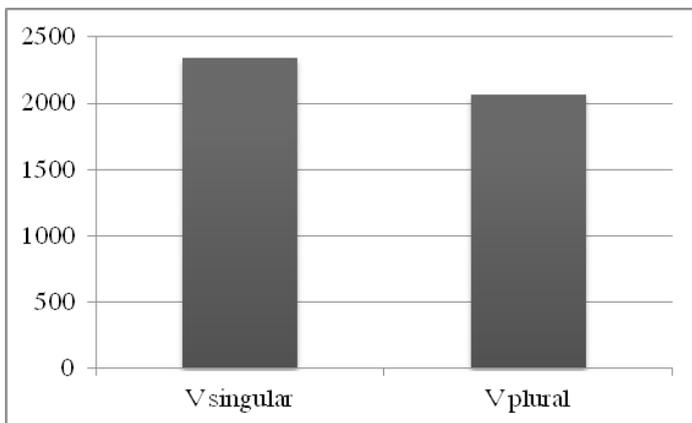
e zero, respectivamente). O gráfico 2 ilustra essa gradação nas médias de tempo de respostas.

Gráfico 2 – Tempo médio de resposta (em ms) em função da distância entre sujeito e verbo



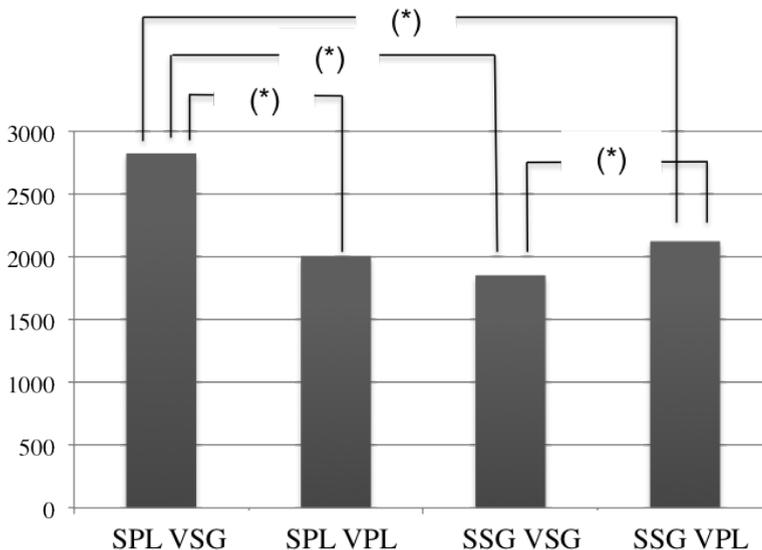
A terceira variável delimitada, qual seja *número do verbo*, também registrou um efeito principal, com médias de tempo de reação significativamente maiores na condição de verbos no singular ($F(1,38) = 8.99, p=0.004$) (Médias: 2338.1ms e 2063.3ms, para verbos no singular e no plural, respectivamente). Como veremos a seguir, esse efeito foi decorrente de um aumento expressivo nos tempos de reação na condição sujeito plural + verbo singular.

Gráfico 3 – Tempo médio de resposta (em ms) em função do número do verbo



Por último, foi obtido um efeito de interação entre *número do sujeito* e *número do verbo* ($F(1,38) = 35.4, p=0.000001$) (Médias: 2822.9ms, 2002.7ms, 1853.2ms e 2123.8ms, para as condições SPL VSG, SPL VPL, SSG VSG e SSG VPL, respectivamente). A condição que registrou as maiores médias de tempo de reação foi a de sujeito plural + verbo singular (Ex. *As músicas... fez...*), justamente a condição que ilustra a concordância verbal variável do PB. Apenas as comparações entre pares que apresentaram resultados estatisticamente significativos são marcadas pelo asterisco no Gráfico 4.

Gráfico 4 – Tempo médio de resposta (em ms) em função da interação entre o número do sujeito e o número do verbo



No que diz respeito ao efeito principal de *número do sujeito* observado, sujeitos plurais parecem facilitar a manutenção na memória da informação relativa a número. Nessas condições foram registrados tempos de reação maiores relativos a leitura/escolha do verbo, quando comparados com os tempos verificados nas condições com sujeitos singulares. Como foi mencionado anteriormente, sujeitos plurais não favorecem a ocorrência de erros de atração (BOCK; EBERHARD, 1993; EBERHARD, 1999; RODRIGUES, 2005a). Vale lembrar que no estudo

de Pearlmutter (2000) só foram observadas diferenças entre distância hierárquica e linear quando o núcleo do sujeito era plural. Além disso, em estudo sobre erros de atração no PB, Rodrigues (2006) não obteve uma produção significativa de lapsos nas condições de sujeito plural. Em conjunto, esses resultados parecem reforçar a ideia de que um DP sujeito plural seria mais saliente e, portanto, seus traços seriam mantidos na memória e recuperados com maior facilidade.

O tempo de reação no segmento crítico foi inversamente proporcional à distância em todas as condições. A distância longa entre sujeito e verbo está associada a menores tempos de reação, independentemente das outras variáveis.

As condições de distância longa favorecem um esvaecimento da informação literal e o que se mantém na memória é uma representação semântica abstrata. No entanto, quando o sujeito é morfofonologicamente marcado, os traços morfológicos do DP sujeito ficam mais acessíveis na memória e, ao localizar o possível verbo da sentença, em caso de incongruência, o participante leva um tempo maior para escolher o verbo do que na condição incongruente em que o sujeito é singular (não marcado). Todavia, em virtude das características da tarefa experimental utilizada, o participante, na ausência de uma alternativa melhor, acaba escolhendo o verbo como opção mais adequada. Esse resultado é compatível com o observado no caso dos lapsos, tanto na produção quanto na compreensão. Como mencionado logo acima, na produção, não há ocorrência de lapsos quando o sujeito é plural (RODRIGUES, 2006). Já na compreensão, Pearlmutter (2000) só registrou efeitos de incongruência quando o sujeito era plural e o verbo singular. Nas condições de sujeito singular, assim como no nosso experimento, não foi registrada diferença.

No caso da condição curta, a representação do DP sujeito com seu conjunto completo de traços morfosintáticos ainda estaria disponível quando da escolha do verbo. Isso aconteceria tanto na condição de sujeito singular quanto de sujeito plural. Assim, a escolha dos verbos nas condições incongruentes registra tempos significativamente maiores do que nas condições congruentes. Na condição de distância zero, a concordância incongruente também apresenta tempos de reação mais altos. Neste caso, contudo, só se revelou estatisticamente significativa a comparação entre SPL VPL e SPL VSG. A comparação entre SSG VSG e SSG VPL, por sua vez, não foi significativa, embora as médias tenham sido na direção esperada (i.e. tempos de reação maiores na condição incongruente).

Diante dos resultados apresentados, pode-se afirmar que tanto a distância quanto a marcação morfofonológica do sujeito parecem ter efeito no processamento da concordância verbal. As médias em cada condição e o resultado da análise estatística da comparação entre os pares de condições relevantes em cada caso (congruente *vs* incongruente), são apresentados na Tabela 2 a seguir.

TABELA 2

Comparações entre pares em função da distância entre sujeito e verbo (médias de tempo de reação em ms por condição e resultado da análise estatística)

Distância entre sujeito e verbo	<i>Comparações entre pares (teste de Tukey)</i>			
	SPL VPL x SPL VSG		SSG VSG x SSG VPL	
Longa	1746.1831ms	2600.35ms	1715.8ms	1854.1ms
	(t(19)=4.14, p= 0.0006)		(t(19)=1.03, p= 0.3150)	
Curta	2109.2ms	2809.4ms	1789.6ms	2121.5ms
	(t(19)=3.62, p= 0.001)		(t(19)=2.43, p= 0.02)	
Zero	2152.9ms	3059.2ms	2054.3ms	2395.9ms
	(t(19)=3.77, p= 0.0013)		(t(19)=1.32, p= 0.2)	

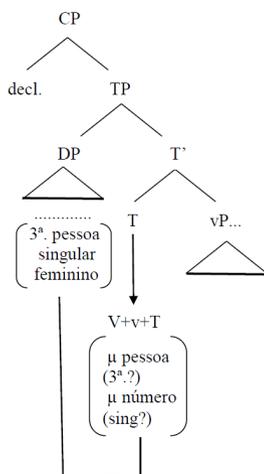
Verbos no plural foram mais rapidamente processados do que os verbos no singular. No entanto, os resultados relativos ao efeito de interação entre *número do sujeito* e *número do verbo* revelaram que a presença de um verbo no singular na condição sujeito plural + verbo singular (*As músicas... fez...*) elevou substancialmente o tempo de reação no segmento crítico. Assim, o efeito principal de *número do verbo* deve ser interpretado levando em consideração o efeito de interação também registrado e discutido a seguir.

A interação entre *número do sujeito* e *número do verbo* mostrou que as condições congruentes (sujeito plural/verbo plural e sujeito singular/verbo singular) não apresentam diferença significativa entre si. Já as condições incongruentes (sujeito plural/verbo singular e sujeito singular/verbo plural) apresentam diferença significativa quando comparadas com as restantes, com tempos de reação maiores – como já foi salientado – para a combinação sujeito plural + verbo singular. Esse

resultado poderia indicar que tal combinação não é percebida como gramatical pelos participantes do teste (universitários, potencialmente falantes da norma culta do PB) e que adotar valores distintos dos da sua gramática para permitir um processamento adequado para a sentença poderia demandar tempo de processamento adicional.

No entanto, o elevado tempo de reação registrado nessa condição também pode ser analisado levando em consideração o que foi discutido na §3 com relação aos mecanismos de recuperação dos traços do sujeito. No processamento, durante a compreensão de enunciados linguísticos, uma operação de checagem de traços é acionada para a computação da concordância entre sujeito-verbo (RODRIGUES *et al.*, 2008). Ocorreria assim, uma verificação dos valores dos traços de número e pessoa entre o sujeito e o verbo.

Figura 2 – Representação formal da operação de checagem de traços na relação sujeito/verbo durante a computação sintática.



Na tarefa experimental conduzida, quando o participante se depara com o verbo da sentença, mecanismos de checagem dos traços devem ser ativados. De acordo com a visão que assume um sistema de *slot* de memória, efeitos de interferência estariam diretamente associados ao esvaecimento de informação na memória de trabalho (cf.

PEARLMUTTER, 2000). Como vimos anteriormente, nos DPs plurais os traços parecem ser menos suscetíveis a esvaecimento. Assim, na condição de sujeito plural do nosso experimento, em que os traços parecem ficar mais salientes, a presença de um verbo no singular parece ser a situação mais conflitante para a computação dos traços de concordância da sentença. Uma incongruência entre os traços de número do sujeito e os do verbo pode ficar mais perceptível nessa configuração do que na outra condição anômala testada: sujeito-singular + verbo-plural. Nesse segundo caso, a incongruência – embora seja detectada pelos participantes, como sugerido pela diferença encontrada entre essa condição (agramatical no PB) e a condição relevante sujeito-singular + verbo-singular (gramatical) – parece gerar um estranhamento menor, possivelmente em virtude de uma subespecificação dos traços do DP sujeito na memória de trabalho.

Conforme visto em § 3, resultados prévios obtidos com crianças e adultos, tanto no que diz respeito à concordância nominal quanto verbal, têm indicado que sentenças contendo concordância não redundante registram tempos de reação significativamente maiores do que sentenças com concordância redundante. Esses resultados incluem experimentos conduzidos com crianças na faixa dos 6-7 anos de idade (o que permite eliminar um eventual papel do ensino formal no julgamento da variedade linguística menos prestigiada socialmente) e alunos cursando o primeiro ano do ensino médio em Centros de Educação de Jovens e Adultos (i.e. falantes cuja gramática potencialmente contemplaria a regra de concordância não redundante) (MARCILESE *et al.*, 2015; AZALIM, 2016; HENRIQUE, 2016; MARCILESE *et al.*, em prep.). Resultados semelhantes (maiores tempos de reação para a concordância não redundante, independente de outras variáveis sociolinguísticas) têm sido reportados também para o inglês (SQUIRES, 2014).

Nesse sentido, nossos resultados vão na mesma direção apontada por esse conjunto de estudos. O fato de as populações testadas nos estudos mencionados serem diversas quanto ao nível de exposição às variantes redundante e não redundante (adultos com escolaridade superior, adultos com ensino fundamental e crianças), aliado à identificação dos mesmos padrões na concordância nominal e verbal, parecem reforçar a análise aqui aventada em termos da relevância da marcação de plural no primeiro item (sujeito da sentença ou determinante, no caso da concordância nominal) na computação dos traços.

5. Breves considerações finais

No presente artigo, buscamos explorar a dimensão processual vinculada a um fenômeno de variação linguística bem conhecido no PB: a concordância verbal. Investigamos especificamente em que medida a distância linear – um fator já estudado no âmbito da psicolinguística experimental – poderia ter um papel relevante no processamento da concordância não redundante por parte de falantes que não necessariamente seriam usuários dessa variante.

Retomando nossas perguntas de pesquisa, formuladas no final de §3, pode-se afirmar que os resultados obtidos são compatíveis com um efeito de distância linear no processamento da concordância, possivelmente ligado a um maior esvaecimento dos traços do sujeito em condições de distância longa entre sujeito e verbo, as quais apresentaram os menores tempos de reação em todas as condições. Os resultados também sugerem diferenças no processamento das regras redundante e não redundante e apontam para a relevância da marcação na atuação dos mecanismos de verificação de traços. Com base no observado nas comparações entre pares, pode-se afirmar que, embora a análise da variância não tenha revelado um efeito de interação entre *distância* e marcação no sujeito e no verbo, esses fatores parecem sim operar de forma conjunta no processamento da concordância.

O tipo de pesquisa aqui desenvolvido – que estabelece um diálogo explícito entre variação e processamento –, bastante exploratória e praticamente inexistente no Brasil, apresenta desafios tanto em termos teóricos quanto metodológicos. A técnica empregada (*maze task*) mostrou-se sensível a efeitos de natureza morfossintática e adequada à investigação de um fenômeno bastante estigmatizado na modalidade escrita culta.

Esperamos com este trabalho ter contribuído para estimular o desenvolvimento de novas pesquisas que visem a estabelecer e estreitar o diálogo entre estudos que atentam para a heterogeneidade do sistema linguístico e abordagens que buscam desvendar a natureza cognitiva da variação.

Referências

ALMEIDA, D. R. *Processamento da concordância de gênero e número em estruturas predicativas*. 2016. 188f. Dissertação (Mestrado). Departamento de Letras. Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2016.

AZALIM, C. *Variação e processamento linguístico: um estudo experimental sobre a concordância nominal variável no PB*. 2016, 129f. Dissertação (Mestrado em Linguística). UFJF, Juiz de Fora, 2016.

AZALIM, C.; MARCILESE, M.; NAME, C.; SCHER, L.; GONÇALVES, L. Concordância nominal variável e saliência fônica: um estudo experimental, submetido.

BOCK, J. K.; CUTTING, J.C. Regulating mental energy: performance units in language production. *Journal of Memory and Language*, Elsevier B.V., v. 31, p. 99-127, 1992.

BOCK, J. K.; EBERHARD, K. M. Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, Taylor & Francis, Routledge, v. 8, p. 57-99, 1993.

BOCK, J. K.; MILLER, C.A. Broken Agreement. *Cognitive Psychology*, Elsevier B.V., v. 23, p. 45-93, 1991.

BRANDÃO, S. F.; VIEIRA, S. R. (Org.). *Ensino de gramática-descrição e uso*. 2. ed. São Paulo: Contexto, 2009.

BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa*, São José Rio Preto, v. 56, n. 3, p. 1035-1064, 2012.

CLAHSEN, H.; HANSEN, D. The missing agreement account of specific language impairment: evidence from therapy experiments. Working Paper. *Essex Research Reports in Linguistics*, University of Essex, Colchester, UK, p. 1-36, 1993.

COSTA, I. de O. *Verbos meteorológicos no plural em orações relativas do Português Brasileiro: sintaxe e processamento*. 2013. 196f. Dissertação (Mestrado) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, 2013.

- COSTA, M. A. As definições de sujeito e seus traços de caracterizadores. O traço de concordância. In: ENCONTRO NACIONAL SOBRE LÍNGUA FALADA E ENSINO, I., 1994, Maceió. *Anais...* Maceió: EDUFAL, 1994. p. 315-320.
- DEUTSCH, A. Subject-predicate agreement in Hebrew: interrelations with semantic processes. *Language and Cognitive Processes*, Taylor & Francis, Routledge, v. 13, n. 5, p. 575-597, 1998.
- EBERHARD, K. The accessibility of conceptual number to the processes of subject-verb agreement in English. *Journal of Memory and Language*, Elsevier B.V., v. 41, p. 560-578, 1999.
- FAYOL, M.; LARGY, P.; LEMAIRE, P. Cognitive Overload and Orthographic Errors: When Cognitive Overload Enhances Subject-Verb Agreement Errors. A Study in French Written Language. *Quarterly Journal of Experimental Psychology*, Taylor & Francis, v. 47, p. 437-467, 1994.
- FORSTER, K. I.; GUERRERA, C.; ELLIOT, L. The maze task: Measuring forced incremental sentence processing time. *Behavioral Research Methods*, Madison, USA, v. 41, p. 163-171, 2009.
- FRANCK, J.; VIGLIOCCO, G.; NICOL, J. Subject-verb agreement in French and English: the role of syntactic hierarchy. *Language and Cognitive Processes*, Taylor & Francis, Routledge, v. 17, n.4, p. 371-404, 2002.
- FREUD, S. A psicopatologia da vida cotidiana. In: SALOMÃO, J. (Org.). *Obras psicológicas completas de Sigmund Freud*: Edição Standard Brasileira, v. VI. Tradução Vera Ribeiro. Rio de Janeiro: Imago, 1996. (Trabalho original publicado em 1901).
- FROMKIN, V. A. The non-anomalous nature of anomalous utterances. *Language*, Linguistic Society of America, v. 47, n.1, p. 27-52, 1971.
- FROMKIN, V. A. *Speech errors as linguistic evidence*. Mouton: The Hague, 1973.
- GARRETT, M. F. The analysis of sentence production. In: BOWER, G. (Ed.). *The psychology of learning and motivation*. New York: Academic Press, 1975. v. 9.

GARRETT, M. F. Syntactic processes in sentence production. In: WALES, R.; WALKER, E. C. T. (Ed.). *New approaches to language mechanisms*. Amsterdam: North-Holland Press, 1976.

GRACIOSA, D. M. D. *Concordância verbal na fala culta carioca*. 1991. 176f. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras, UFRJ, Rio de Janeiro, 1991.

HENRIQUE, K. *Variação linguística e processamento: investigando o papel da distância entre sujeito e verbo na realização da concordância verbal variável no PB*. 2016, 149f. Dissertação (Mestrado) – Faculdade de Letras, Universidade Federal de Juiz de Fora, Juiz de Fora, 2016.

NETO, Nelson Lima. Contas vazias do Estado deixa pagamento de servidores, pela Justiça, sem prazo para acabar. *Extra Digital*, Rio de Janeiro, 13 set. 2016. Disponível em: <<http://extra.globo.com/emprego/servidor-publico/contas-vazias-do-estado-deixa-pagamento-de-servidores-pela-justica-sem-prazo-para-acabar-20098835.html#ixzz4bxdqp9TE>>. Acesso em: 13 set. 2016.

LEMLE, M.; NARO, A J. *Competências básicas do Português*. Rio de Janeiro: Mobral/Fundação Ford, 1977.

LOPES, N.; BAXTER, A. N. A concordância verbal variável no Português dos Tongas. *Papia*, Brasília, v. 21, p. 39-50, 2010.

MARCILESE, M; HENRIQUE, K.; AZALIM, C.; NAME, C. Processamento da concordância variável no PB em uma perspectiva experimental. *Linguística*, Rio de Janeiro, UFRJ, v. 11, n.1, p.118-134, 2015.

MARCILESE, M.; NAME, C.; LAIER, P.; MAIA, M.; SCHER, L.; GONÇALVEZ, L.; AZALIM, C. Aquisição da linguagem e variação linguística: um estudo experimental sobre a concordância nominal variável na fala infantil. Em preparação.

MENDES, R. B.; OUSHIRO, L. Variable number agreement in Brazilian Portuguese: an overview. *Language and Linguistics Compass*, John Wiley & Sons, Inc, v. 9, n. 9, p.358-368, 2015.

MERINGER, R.; MEYER, K. *Versprechen und Verlesen*. Stuttgart: [s.n.], 1885.

MITCHELL, D. C. On-line methods in language processing: Introduction and historical review. In: CARRERAS, M.; CLIFTON, C. *The on-line study of sentence comprehension: eyetracking, ERPs, and beyond*. New York: Psychology Press, 2004.

MOTTA, M. F. *O fenômeno variável da concordância verbal em redações da EJA: uma abordagem sociolinguística*. 2011. 135 f. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

NARO, A. The social and structural dimensions of a syntactic change. *Language*, Linguistic Society of America, v.57, n.1, p.63-98, 1981.

NICOLAU, E. M. D. *A ausência de concordância verbal em português: uma abordagem sociolinguística*. 1984. 152f. Dissertação (Mestrado em Linguística) – Universidade Federal de Minas Gerais, Belo Horizonte, 1984.

PEARLMUTTER, N.J. Linear versus hierarchical agreement feature processing in comprehension. *Journal of Psycholinguistic Research*, Springer Nature, v. 29, n. 1, p. 89-98, 2000.

PHILLIPS, C.; WAGERS, M. W.; LAU, E.F. Grammatical illusions and selective fallibility. In: RUNNER, J. (Ed). *Experiments at the interface*. Syntax & Semantics. Bingley, UK: Emerald Publications, 2011. v. 37 p. 153-186.

RODRIGUES, E. dos S. O processamento da concordância verbal com construções partitivas no português brasileiro. *Linguística*, Rio de Janeiro, UFRJ, v. 1, n.1, p. 147-170, 2005a.

RODRIGUES, E. dos S. Distinguindo aspectos sintáticos, pré-sintáticos e pós-sintáticos no processamento da concordância gramatical. In: MIRANDA, N. S.; NAME, M. C. L. (Org.). *Linguística e cognição*. Juiz de Fora: Editora da UFJF, 2005b.

RODRIGUES, E. dos S. *Processamento da concordância de número entre sujeito e verbo na produção de sentenças*. 2006. 202f. Tese (Doutorado) – Departamento de Letras – PUC-Rio, Rio de Janeiro, 2006.

RODRIGUES, E. dos S. Concordância verbal com construções partitivas: uma proposta de análise. *Veredas*, Juiz de Fora, UFJF, v. 1, p. 93-107, 2011.

RODRIGUES, E. dos S.; CORREA, L. M. S.; AUGUSTO, M. R. A. Concordância sujeito-verbo em um modelo integrado misto (top-down/bottom-up) da computação on-line. *Veredas*, Juiz de Fora, UFJF, v. 2, p. 76-91, 2008.

RUBIO, C. F. *Padrões de concordância verbal e de alternância pronominal no português brasileiro e europeu: estudo sociolinguístico comparativo*. 2012. 391 f. Tese (Doutorado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2012.

RUBIO, C.F. Concordância verbal de terceira pessoa do plural no português europeu: variação ou regra semicategórica? *Fórum Linguístico*, Florianópolis, v.12, n.3, p.786-806, 2015.

SANTOS, R. L. de A. *A concordância verbal na fala de menores carentes que vivem em entidades filantrópicas de Maceió*. 2010. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal de Alagoas, Maceió, 2010.

SCHERRE, M. M. P.; NARO, A. J. Duas dimensões do paralelismo formal na concordância de número no português popular do Brasil. *D.E.L.T.A. – Revista de Documentação de Estudos em Linguística Teórica e Aplicada*, São Paulo, EDUC, v. 9, n. 1, p. 1-14, 1993.

SCHERRE, M. M. P.; NARO, A. J. A concordância de número no português do Brasil: um caso típico de variação inerente. In: HORA, D. da (Org.). *Diversidade linguística no Brasil*. João Pessoa: Ideia, 1997. p. 93-114.

SCHERRE, M. M. P.; NARO, A. J. Sobre a concordância de número no português falado do Brasil. In: RUFFINO, G. (Org.). *Dialettologia, geolinguística, sociolinguística*. Centro di Studi Filologici e Linguistici Sicilliani. Università di Palermo. Tübingen: Max Niemayer Verlag, 1998a. p. 509-523.

SCHERRE, M. M. P.; NARO, A. J. Restrições sintáticas e semânticas no controle da concordância verbal em português. *Fórum Linguístico*. Florianópolis, n. 1, p.45-71, p. 54-71, 1998b.

SILVA, J.A.A. *A concordância verbal de terceira pessoa do plural do Brasil: um panorama sociolinguístico de três comunidades no interior do Estado da Bahia*. 2005. 323f. Tese (Doutorado) – Programa de Pós-Graduação em Letra e Linguística da UFBA, Salvador, 2005.

SQUIRES, L. Social differences in the processing of grammatical variation. *Selected Papers from NWAV 42*, University of Pennsylvania, v. 20, n.2, p. 178-188, 2014.

VIEIRA, S. (Org.). *A concordância verbal em variedades do português*. A interface fonética-morfossintaxe. Rio de Janeiro: Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro, 2015.

VIEIRA, S. R. Aspectos da concordância verbal em dialetos populares. In: ENCONTRO NACIONAL SOBRE LÍNGUA FALADA E ENSINO, I., 1994, Maceió. *Anais...* Maceió: EDUFAL, 1994. p. 323-327.

VIEIRA, S. R. Aspectos da concordância verbal em dialetos populares. In: ENCONTRO NACIONAL SOBRE LÍNGUA FALADA E ENSINO, I., 1994, Maceió. *Anais...* Maceió: EDUFAL, 1994. p. 323-327.

VIGLIOCCO, G.; NICOL, J. The role of syntactic tree structure in the construction of subject-verb agreement. Unpublished manuscript. *Linear vs Hierarchical Agreement* 11, 1997.

VIGLIOCCO, G.; NICOL, J. Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition*, Elsevier Science B. V., v. 68, B13-B29, 1998.

VIGLIOCCO, G.; BUTTERWORTH, B.; GARRETT, M. F. Subject-verb agreement in Spanish and English: Differences in the role of conceptual constraints. *Cognition*, Elsevier Science B. V., v. 61, p. 261-298, 1996.

VOTRE, S. Relevância da variável escolaridade. In: MOLLICA, Maria Cecilia; BRAGA, Maria Luiza (Org.). *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2013.

WAGERS, L.; LAU, E. F.; PHILLIPS, C. Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language*, Elsevier B. V., v. 61, p. 106-237, 2009.

Agreement effects of gender and number in pronominal coreference processing in Brazilian Portuguese

Efeitos de concordância de gênero e número no processamento da correferência pronominal em português brasileiro

Michele Calil dos Santos Alves

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil

michelecalil@gmail.com

Abstract: Pronominal coreference is a syntactic dependency in which pronouns are bound to previous referents in discourse. One of the keys to understanding coreference processing is memory, since information that has already been interpreted and stored must be integrated with new material in real time. The aim of this research is to investigate how pronominal antecedents are retrieved from memory, and more precisely to clarify the role of structural constraints, agreement features and decay factors. Since Brazilian Portuguese has rich morphology, speakers of this language can rely on agreement cues as well as structural constraints to resolve coreference. The hypothesis is that candidates that feature-match pronouns will initially influence coreference processing, even though they violate structural constraints, which will only work later in binding processing to help the parser select the most adequate antecedent. An eye-tracking experiment was conducted with twenty-four native speakers of Brazilian Portuguese. The results showed that structurally unacceptable antecedent candidates that feature-matched the pronouns in gender and number facilitated coreference processing. It is claimed that they were considered as potential antecedents. Moreover, it seems that memory might be sensitive to differences that exist between singular and plural features. Plural may be more salient in memory due to the fact it is marked

in Brazilian Portuguese. Finally, memory can also be affected by decay factors, which, for example, can be responsible for processing difficulties when there is a long distance between antecedents and pronouns.

Keywords: coreference processing; gender and number features; Brazilian Portuguese.

Resumo: A correferência é uma dependência sintática em que pronomes são ligados a antecedentes prévios no discurso. Uma das chaves para se compreender o processamento da correferência é a memória, uma vez que as informações que foram anteriormente interpretadas e armazenadas devem ser integradas com novo material em tempo real. O objetivo desta pesquisa é investigar como os pronomes recuperam seus antecedentes na memória, mais precisamente, esclarecer o papel das restrições estruturais, dos traços de concordância e dos fatores de *decay*. Uma vez que o português brasileiro possui rica morfologia, leitores desta língua podem utilizar-se tanto de traços de concordância, como de restrições estruturais para resolver a correferência. A hipótese era de que candidatos que concordassem com os pronomes em gênero e número poderiam influenciar o início do processamento da correferência, apesar de violarem as restrições estruturais, e estas, por sua vez, somente entrariam em jogo mais tarde para ajudar o processador a selecionar o antecedente mais adequado. Foi realizado um experimento de monitoramento ocular com vinte e quatro falantes nativos de português brasileiro e os resultados apontam que antecedentes estruturalmente inaceitáveis que concordam com os pronomes em gênero e número facilitaram o processamento da correferência. Eles foram considerados antecedentes em potencial, apesar de violarem as restrições estruturais. Além disso, parece que a memória pode ser sensível às diferenças existentes entre singular e plural. O plural pode ser mais saliente na memória devido ao fato de ser marcado em português brasileiro. Finalmente, a memória também pode ser afetada por fatores *decay*, que, por exemplo, podem ser responsáveis por dificuldades de processamento quando há uma distância longa entre os antecedentes e os pronomes.

Palavras-chave: processamento da correferência; traços de gênero e número; português brasileiro.

Recebido em: 11 de dezembro de 2016

Aprovado em: 7 de maio de 2017

1. Introduction

In order to process language in real time, previously interpreted information must be kept at least momentarily in our memory so that integration with novel upcoming material can take place rapidly (LEWIS *et al.*, 2006). In this way, memory can be considered one of the key factors in processing long distance dependencies such as coreference, in which pronouns are bound to antecedents that occupy linearly distant positions in the discourse.

Among other cues, binding can be influenced by structural constraints, agreement relations between antecedents and anaphors, and salience of the discourse entities involved in the context. Previous research that has investigated how those three factors play a role in binding processing is very contradictory. On the one hand, it has been claimed that structurally unacceptable candidates cannot initially influence binding processing even in cases in which they are salient discourse entities and agree with the anaphors (NICOL; SWINNEY, 1989; CLIFTON *et al.*, 1997; STURT, 2003; LEITÃO *et al.*, 2008; XIANG *et al.*, 2009; OLIVEIRA *et al.*, 2012; DILLON *et al.*, 2013; CHOW *et al.*, 2014). On the other hand, other research has shown that structural constraints can be fallible as these studies found that structurally unacceptable candidates can be initially considered as potential antecedents if they are salient entities that feature-match the anaphors (BADECKER; STRAUB, 2002; KENNISON, 2003; PARKER, 2014; PATIL *et al.*, 2016).

One possible explanation for these contradictory results in relation to the role of agreement in binding processing in the literature may rely on the fact that those studies may have taken for granted intrinsic differences that exist among morphological features. In these terms, our research tried to control for the different types of features that may exist under the category of gender (masculine and feminine) and number (singular and plural). In addition, English may not be the most appropriate language to study agreement, as it is a language with limited overt morphology. By comparing overt agreement marking in English and in Brazilian Portuguese, one notices that unlike the former, the latter has redundant agreement marking in most determiners, nouns, adjectives, and verbs, for example. Sentence (1) shows how one of the sentences used in one of our experiments in Brazilian Portuguese would be translated into English. The sentence in Brazilian Portuguese (1a) has 17 overt marks, while its translation in English (1b) has only 8.

- (1) a) O[masc, sg] engenheiro[masc, sg] investigou[sg] os[masc, pl] arquitetos[masc, pl] que roubam[3rd person, pl] ele[masc, 3rd person, sg] há alguns[masc, pl] semestres[pl].
- b) The engineer[sg] investigated the architects[pl] who have[3rd person, pl] stolen him[masc, 3rd person, sg] for a couple of semesters[pl].

Lago *et al.* (2015) compared agreement attraction in subject-verb dependencies in Spanish (another morphologically rich language similar to Brazilian Portuguese), and in English. Their results showed that Spanish comprehenders showed more processing difficulties in ungrammatical sentences than English comprehenders. Moreover, Spanish comprehenders, but not English comprehenders showed processing difficulties in grammatical sentences with plural attractors.¹ The authors explain that since agreement morphology is functionally

¹ Sample of the materials of Lago et al (2015)

Experiment in Spanish:

Gram, sg attractor: La nota que la chica **va** a escribir en la clase alegrará a su amiga.

(The note that the girl are going to write during class will cheer her friend up.)

Gram, pl attractor: Las notas que la chica **va** a escribir en la clase alegrará a su amiga.

(The notes that the girl are going to write during class will cheer her friend up.)

Ungram, sg attractor: *La nota que la chica **van** a escribir en la clase alegrará a su amiga.

(The note that the girl are going to write during class will cheer her friend up.)

Ungram, pl attractor: *Las notas que la chica **van** a escribir en la clase alegrará a su amiga.

(The notes that the girl are going to write during class will cheer her friend up.)

Experiment in English:

Gram, sg attractor: The musician that the reviewer **was** highly praising last week will probably win a Grammy.

Gram, pl attractor: The musicians that the reviewer **was** highly praising last week will probably win a Grammy.

Ungram, sg attractor: *The musician that the reviewer **were** highly praising last week will probably win a Grammy.

Ungram, pl attractor: *The musicians that the reviewer **were** highly praising last week will probably win a Grammy.

more important in Spanish than in English, Spanish speakers would rely more on morphological cues in processing sentences. Therefore, the strength of agreement predictions would be higher for Spanish than in English, which causes a higher pay off when the predictions are not fulfilled and reanalysis is needed.

Taking the fact that the use of agreement cues may be more fruitful in languages with rich morphology like Spanish, the present work aims to investigate how pronouns retrieve antecedents in Brazilian Portuguese, which is also a language with rich morphology. Moreover, it seems that the use of morphological cues in memory retrieval may also vary depending on the particular binding dependency. Agreement features may be more helpful in pronominal antecedent retrieval due to the looseness of Principle B, since it only posits that the pronoun antecedent must not be local [see section 3 of the present paper].

The recognition of a pronoun must initiate a retrospective search for an antecedent. Since the structural relation between a pronoun and its antecedent is almost free, it is natural to assume that a pronoun initiates a cue-based search for an antecedent that shares its person, number, and gender features, and hence it would not be surprising for this search to detect nouns that match those cues, even when they violate Principle B (PHILLIPS; WAGERS; LAU, 2011, p. 171)

In this way, the present research will fill a gap in the literature as it will provide not only one more piece of evidence to the puzzle of binding processing, which lacks intensive investigation, but it will determine whether different types of gender and number features carried by candidates are responsible for differences in the way memory retrieves the antecedents. It will also be determined whether speakers of languages with morphological richness such as Brazilian Portuguese tend to initially consider structurally unacceptable candidates as potential antecedents despite the fact that they violate binding constraints. In addition, pronouns will be our object of the study since they might rely more on content cues as opposed to reflexives.

Thus the main aim of this research is to investigate how pronouns retrieve antecedents in Brazilian Portuguese. In addition, its secondary aim is to examine which features can influence memory retrieval the most. The first hypothesis is that candidates that feature-match the pronouns would initially influence coreference processing in Brazilian Portuguese, even though they violate Principle B, and that structural constraints

would only work later on in binding processing to help the parser select the most appropriate antecedent (cf. BADECKER; STRAUB, 2002).

In addition, it is hypothesized that memory and consequently coreference processing is sensitive to different types of agreement features such as masculine and feminine, for gender; and singular and plural, for number. Since feminine and plural are marked features in Brazilian Portuguese, we expect that they will be more salient in memory, making the antecedent candidates that carry these types of features more easily retrieved. The correlation between the influence of the structurally unacceptable antecedents and the types of features they display is known as the *mismatch asymmetry*. It seems that structurally unacceptable candidates with marked features are more influential than those with unmarked features (cf. among others for plural and singular, BOCK; MILLER, 1991; WAGERS *et al.*, 2009; DILLON, 2013).

Finally, it is also hypothesized that memory is affected by decay over time (LEWIS; VASISHTH, 2005; LEWIS; VASISHTH; VAN DYKE, 2006), so that a long linear distance between antecedent candidates and anaphors brings costs to binding processing.

In order to test the hypotheses, an eye-tracking experiment was conducted with native speakers of Brazilian Portuguese. The eye-tracking technique is suitable for our purposes as it enables the researcher to examine the temporal course of language processing, including early (First Fixation Duration²) and late (Total Fixation Duration³) on-line processing measures.

This paper will be arranged as follows: Section 2 will present the reader with the computational model that is commonly used in the literature to explain how memory retrieval operates; in Section 3 the structural constraints on coreference called the Binding Principles will be briefly reviewed; Section 4 will discuss previous research in the literature; Section 5 will introduce the present study; Section 6 will discuss the main conclusions of this study, followed by References.

² First Fixation Duration is the duration of the first fixation in a certain text region.

³ Total Fixation Duration is the sum of all the fixations in a certain text region.

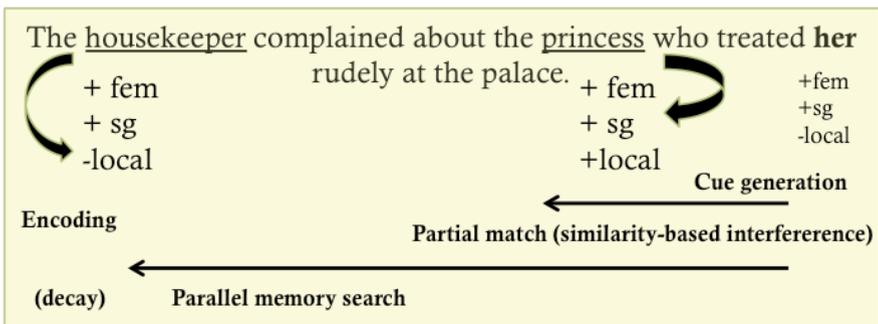
2. Content-Addressable Memory (CAM)

Content-addressable memory (CAM) (McELREE, 2000; McELREE *et al.*, 2003; van DYKE; McELREE, 2006) is a computational model that has been used recently to explain how memory operates during language processing. In this model, previous information that was previously interpreted can be retrieved by a parallel search based on a set of grammatical cues generated by the target. This parallel search in memory can be affected by similarity-based interference and decay factors (LEWIS; VASISHTH, 2005; LEWIS; VASISHTH; van DYKE, 2006). The former occurs when the similarity between items in memory and the retrieval cues increase, reducing the strength of association between the cue and the target item, as a great number of items will be associated with the cue. Consequently, memory failure rates increase, and distractors, that is, candidates that partially-match the cues can sometimes be retrieved. The latter occurs when the linear distance between the dependent items is increased and the activation of the distant item decays over time, which makes its retrieval more difficult.

Retrieval cues consist of several types, including structural, morphological, semantic, and contextual cues (among others). The present paper will focus on only two of them: structural and morphological cues.

Figure 1 (based on LEWIS; VASISHTH; van Dyke, 2006) illustrates how pronouns retrieve their antecedents in memory.

FIGURE 1 – How antecedent retrieval works in CAM. Figure based on Lewis, Vasishth and van Dyke (2006)



During the encoding phase, all information is interpreted and stored in memory. By the time the pronoun is encountered, a group of grammatical cues is generated in order to retrieve the antecedent. In the example portrayed in Figure 1, the antecedent must not be local,⁴ and it must be feminine and singular. After that, there is a parallel search in memory and two candidates that are similar to the cues generated by the target are found: “housekeeper” and “princess”. The former candidate is a perfect match; however, although the latter candidate is only a partial match as it is local, it can interfere with memory retrieval, the so-called similarity-based interference effect. Candidates like “princess” are called distractors according to the CAM model. In addition, in this example, “housekeeper” can also decay over time as it was stored in memory before “princess”, which, in this case, is more recent. Thus, according to this model, distractors such as “princess” can sometimes be erroneously retrieved as antecedents as a result of a failure caused by both similarity-based interference effects and decay factors.

3. Binding Principles

The Binding Theory (CHOMSKY, 1993) posits three principles: A, B, and C, which are able to explain, respectively, the distributional constraints on (a) anaphor (according to Chomsky, only includes the reflexives and reciprocals); (b) pronouns; and (c) free referential expressions.

Chomsky (1993) claimed that depending on the nature of the NPs involved and the syntactic configurations in which they occur, the anaphoric relations can be possible, necessary, or proscribed.

(2) John said Mary criticized him.

(3) John criticized him.

Chomsky (1993) states that in (2), *him* can take *John* as its referent, which cannot happen in (3). According to the Binding Principle B, pronouns cannot have a structurally local antecedent. It is noteworthy that

⁴ It is important to mention that the status of the [-local] feature can be questioned, as it seems awkward that languages would have this feature specified for each and every item. However, we assume that it is actually a relational feature that is only specified in binding dependencies.

this is not a matter of linear distance, as pronouns can actually linearly precede their antecedents, like in (4), a construction traditionally known as *cataphora*. Moreover, in (5), *him* cannot refer to *John*, even though there is a long linear distance between them.

- (4) After he entered the room, John sat down.
- (5) He said Mary criticized John.

Thus Chomsky proposed that a pronoun couldn't take as its antecedent an element within its [binding] domain. In (3), the domain of the pronoun is the whole sentence; therefore, as *John* is within the domain of *him*, it cannot be its referent. On the other hand, in (2) and (4), *John* is not in the same domain of *him* since they are in different clauses.

- (6) α binds β if α c-commands β , and α and β are co-indexed.
- (7) If β is not bound, β is free.
- (8) A referential expression (neither a pronoun nor an anaphora) must be free

The previous sentences with their indexes are the following:

- (9) John_i said Mary criticized him_i.
- (10) After he_i entered the room, John_i sat down.
- (11) *He_i said Mary criticized John_j.
- (12) He_i said Mary criticized John_j.

The example (11) can only be grammatical if *him* and *John* have different indexes, like in (12).

- (13) If the α index is different from the β index, α cannot be the antecedent of β and vice-versa.

The example in (3) with indexes would be like:

- (14) John_i criticized him_j.
- (15) * John_i criticized him_i.

By comparing (9) and (15), one notices that a pronoun is able to exist within the binding domain of its antecedent. However, it should be highlighted that a pronoun cannot be too close to its antecedent.

(16) A pronoun must be free in a local domain (Principle B)

The local domain is generally the minimum clause that contains the pronoun. Unlike pronouns, which *can* have a binding antecedent but do not *need* one, anaphora (reflexives and reciprocals), like in (17), require antecedents to bind them. In addition, the antecedents of anaphora need to be in the same local domain:

(17) John_i criticized himself_i.

(18) An anaphor must be bound in its local domain (Principle A).

Clearly, pronouns cannot be substituted for anaphora:

(19) *John_i said Mary criticized himself_i.

(20) *John_i said Mary criticized himself_i.

Finally, Chomsky (1993) postulates the Binding Principles as the following:

(21) Principle A: an anaphor must be bound in its local domain.

Principle B: a pronoun must be free in its local domain.

Principle C: an R-expression must be free.

Principles A and B had as their local domain the minimum clause that contains the anaphor or pronoun. However, the rule of the minimum domain does not address (22) and (23):

(22) John_i believes [himself_i to be clever].

(23) *John_i believes [him_i to be clever].

In (22) *himself* is not bound in its local domain, but it is a grammatical sentence, whereas in (23) *him* is free in its local domain, but it is an ungrammatical sentence. The answer to this can be found in

government. In the examples (22) and (23), the main verb *believe* governs *himself* and *him* due to the fact that *to be* does not carry inflection. The concept of local domain must therefore be substituted with governing category. This way, the governing category is (22) and in (23) is the whole sentence, and Principles A and B are no longer violated. Chomsky (1993) explains that the governing category of α is the minimum Complete Functional Complex that contains α and in which the binding principle of α can be satisfied.

- (24) Principle A: an anaphor must be bound in its governing category
Principle B: a pronoun must be free in its governing category.
Principle C: an R-expression must be free.

Psycholinguists have been trying to investigate whether the principles of the Binding Theory would rapidly influence binding processing during on-line sentence processing. In the next section some studies that examined the influence of binding principles in online language processing will be reviewed.

4. Previous research in binding processing

In this section, some previous research on coreferential processing with respect to Principles A and B will be reviewed⁵. The relationship between these structural constraints and agreement cues in the time-course of binding processing is very controversial in the literature. Therefore, previous research will be presented here under two subsections: works that showed some evidence of initial infallibility of structural constraints in binding processing; and works that found the opposite, that is, structurally unacceptable candidates can be initially considered as potential candidates if they feature-match the anaphoric expressions.

⁵ Principle C processing will not be addressed here since, unlike Principles A and B, it does not have any structural constraints operating on it.

4.1 Evidences of initial infallibility of structural constraints in binding processing

Nicol and Swinney (1989) conducted a cross-modal priming experiment examining the reactivation of anaphoric antecedents. They found out that immediately after the anaphor only the structurally appropriate antecedent was reactivated, while the other referents were not significantly reactivated. The results for pronouns were similar to the results for anaphora. Thus, the authors concluded that the reactivation of prior referents is restricted by grammatical constraints. Nicol & Swinney (1989) explained that only when binding constraints do not constrain the list of potential antecedents to a single one; pragmatic and other sentence or discourse processing procedures would come into play, but only at a later point in processing.

Clifton *et al.* (1997) studied how antecedents of “her” and “him/his” are reactivated. They performed a phrase-by-phrase self-paced moving window experiment contrasting noun phrase (NP) and specifier (SPEC) usages. They also manipulated the morphological number of the subject in each sentence. The authors found faster reading times for the SPEC trials when the number of the subject agreed with the pronoun, which would make it an appropriate antecedent. However, when the subject and the pronouns mismatched in number, there was a slowdown on reading times as the subject was made inappropriate. Importantly, number did not show any effects on NP trials. Thus Clifton and colleagues concluded that, at least initially, binding principles constrain parsing decisions, and that number would work as a filter to determine whether the accessed antecedents are appropriate.

Sturt (2003) was concerned about two questions: i) to what extent sentence processing is affected by ungrammatical antecedents; ii) to what extent do binding principles act like a filter on the final interpretation of a sentence. He conducted an eye-tracking study to investigate the influence of inaccessible antecedents in reflexive binding when they are put strongly into discourse focus. Stereotypical subjects were used in order not to expose participants to ungrammatical sentences. His results show that binding constraints were applied extremely early (at First Fixation and First Pass reading times). First Fixation and First Pass reading times were faster when the gender of the reflexive matched the stereotype of the accessible antecedent than it did not, but they did not differ reliably as a function of whether the inaccessible antecedent matched the reflexive.

However, reliable influences of the inaccessible antecedent at late measures were found (Second Pass in the second area after the reflexive). There were longer Second Pass times when the inaccessible antecedent mismatched the reflexive than when it did not. The author concluded that antecedents that were not initially considered by the binding principles could affect processing at a later stage. In other words, binding constraints are applied at an extremely early stage, but they do not act as filters. Sturt (2003) also conducted a follow-up study, a sentence-by-sentence self-paced reading experiment with a comprehension question to check the interpretation of the anaphor referent. It seems that Principle A did not act as an absolute filter on the final interpretation of the sentence either. Sturt (2003) defends the idea that binding principles act like a defeasible filter, as they can be violated at a later stage when there is a highly focused unacceptable antecedent involved.

Leitão *et al.* (2008) investigated the relationship between Principle B and phi-features (gender, number, and animacy) in coreference processing in Brazilian Portuguese in two self-paced reading experiments. In the first experiment, there were structurally unacceptable antecedents in the sentences, and the results showed that the pronoun+1 region (adverb regions) had longer reading times due to the fact that the structurally unacceptable antecedent in the sentence feature-matched the pronoun. However, in the second experiment, there was a structurally unacceptable candidate available in a preamble. Unlike the first experiment, the results of the second experiment did not show any differences among the conditions, although the reading times at the pronoun region were faster when compared to the first experiment. The authors suggested that when there are no structurally acceptable antecedent candidates available, as in the first experiment, candidates that feature-match the pronouns could be considered as potential antecedents even if they violate Principle B. On the other hand, when there is a structurally acceptable antecedent available, as in the second experiment, the search of an antecedent ends faster and the structurally unacceptable candidates are not taken into account.

In an *Event Related Potentials* experiment (ERPs), Xiang *et al.* (2009) studied intrusion effects of structurally unacceptable noun phrases that matched the reflexive. The authors found a P600-like component for both intrusive and incongruent conditions. However, there were no differences between the intrusive and incongruent sentences, while both were significantly different from the congruent. It is important that they

found a marginal late intrusion effect only at 800-1000ms, which matches the late effects of inaccessible antecedents in Sturt (2003). The authors concluded that there is no initial intrusion effect for reflexive binding.

Oliveira *et al.* (2012) conducted a self-paced reading experiment to determine whether Principle A influences reflexive resolution in Brazilian Portuguese. They found that the grammatical conditions, in which the structurally acceptable antecedent agrees in gender with the reflexives, had faster reading times at the reflexive region when compared to ungrammatical conditions. It should be noted that the structurally unacceptable antecedents were not taken into account in any condition, which suggests that Principle A works as a filter, blocking the candidates that violate it.

Dillon *et al.* (2013) conducted eye-tracking experiments with the purpose of investigating the impact of structurally illicit nouns phrases on the computation of reflexive binding. It should be mentioned that they also conducted off-line judgments to check whether the number mismatch in the materials would be reliably rejected. The results of the offline grammaticality judgment indicated a main effect of grammaticality, confirming this. Likewise, the online results showed a main effect of grammaticality in First Pass and in Total Times, with no facilitatory intrusion effects. The authors concluded that initially the feature content of a structurally illicit NP could not affect reflexive processing. Thus they concluded that the mechanism used by memory retrieval for reflexives primarily uses syntactic information to guide retrieval of the antecedents. It is relevant to the current study that comprehenders seemed to be less sensitive to the feature match when the head noun was plural suggests that the feature mismatch is sensitive to the markedness of the features involved.

Chow *et al.* (2014) were concerned about which kinds of constraints initially restrict antecedent retrieval, and which have later effects, working as filters. In their first self-paced moving window experiment they manipulated the gender match between the pronoun “him” and the structurally acceptable main clause subject and the structurally unacceptable embedded clause subject. Relative clauses could also modify the nouns in order to increase the distance between the pronoun and the antecedent. The structurally unacceptable antecedents could be either a common noun or a proper name. As the mismatch conditions had longer reading times, it seems that comprehenders are immediately

sensitive to the structural constraints on pronoun interpretation regardless of the similarity between the candidate antecedents and linear distance. They found robust effects of grammaticality, but no interference effects of any kind. It should be mentioned that when the linear distance between the pronoun and the structurally acceptable antecedent was long in the modified common noun condition, they found a late ungrammatical match effect, that is, when no grammatical antecedent was available, the presence of a feature-matching structurally unacceptable antecedent led to longer reading times. The authors explain that it may have been caused by the fact that the memory representation of the structurally acceptable antecedent was decayed due to the long distance. In their second experiment, Chow *et al.* (2014) tried to replicate the results found on Badecker & Straub (2002) [which will be discussed in the next subsection] by using identical materials and procedures. However, Chow *et al.* (2014) failed and only replicated the results of their first experiment. They also conducted 3 other experiments, but no effects were found. The authors defended the *Simultaneous Constraints hypothesis* since it appeared that both agreement features like gender and the structural constraints of binding immediately restricted the set of candidate antecedents during the initial retrieval process.

4.2 Evidence of the initial fallibility of the structural constraints in binding processing

Badecker and Straub (2002) studied the processing of reflexive and pronoun binding in a series of self-paced reading experiments. According to the authors, coreference processing is influenced by: morphological and syntactic properties of the dependent expression and the antecedents; structural parallelism; causal semantics; prominence and salience of the local discourse entities; and the world knowledge shared about the discourse entities involved. Among these factors, the authors' study was focused on morphosyntactic features and local focus of attention. In one of their experiments, they investigated whether the content of structurally inaccessible NPs would influence pronoun processing.

- (25) a) multiple match: John thought that Bill owed **him** another chance to solve the problem.
- b) accessible match: John thought that Beth owed **him** another chance to solve the problem.
- c) inaccessible match: Jane thought that Bill owed **him** another chance to solve the problem.
- d) no-match: John thought that Beth owed **him** another chance to solve the problem.

They observed longer reading times in the no-match condition than in the accessible match condition. The results also show faster reading times when there was a structurally accessible antecedent than when there was an inaccessible antecedent. There was no difference between the multiple match and the accessible-match conditions. The authors concluded that gender was automatically used to identify the referent of a pronoun, and that the structurally accessible antecedents were also rapidly accessed. On the other hand, inaccessible candidates were not blocked for an initial candidate set, as they influenced the evaluation process as soon as the pronoun was encountered.

Badecker and Straub (2002) also investigated whether number features could shape the initial candidate set. In another experiment, they studied the influence of grammatical number in reciprocal anaphors like “each other”, which are also governed by Principle A, as can be seen in (26):

- (26) a) multiple match: The attorney thought that the judges were telling **each other** which defendants has appeared as witnesses before.
- b) single-match: The attorneys thought that the judges were telling **each other** which defendants has appeared as witnesses before.

The results indicate longer reading times in the multiple-match than in the single match, but only 3-4 words after the anaphor. The authors suggested that morphological number contributes to identifying the initial set of antecedent candidates. The multiple-match effect was attenuated

in this case, because, according to the authors, common nouns may not be as effective as proper names in establishing discourse entities.

Badecker and Straub (2002) concluded that binding-theory principles do not function as initial filters as reading times were longer when the grammatically inaccessible NPs agreed in gender (and number) with the pronoun or anaphor. The authors supported the *interactive-parallel-constraint model*: the initial candidate set is composed of the focused discourse entities that are compatible with the lexical properties of the referentially dependent expression, while the grammatical constraints on interpretation operate quickly and effectively in the process of selecting from among these options.

Kennison (2003) investigated how comprehenders use structural information during coreference resolution of the pronouns “her”, “him”, and “his”. In a self-paced moving window experiment, Kennison (2003) examined the processing of “her” in object position, functioning as either an NP or SPEC as in (27).

(27) SPEC conditions:

Susan watched **her** classmate during the open rehearsals of the school play.

Carl watched **her** classmate during the open rehearsals of the school play.

They watched **her** classmate during the open rehearsals of the school play.

NP conditions:

Susan watched **her** during the open rehearsals of the school play.

Carl watched **her** during the open rehearsals of the school play.

They watched **her** during the open rehearsals of the school play.

She found that the type of subject influenced coreference processing in both conditions, including in NP conditions, which is inconsistent with Nicol and Swinney (1989) and Clifton *et al.* (1997). In SPEC conditions, reading times were longer when the subject was a male name, while in NP conditions reading times were longer when the subject was female. And the shortest times were for the conditions

with “they”. In other words, when coreference could be achieved, there were longer reading times for NP conditions than for SPEC, as SPEC conditions were easy to process. However, when coreference could not be achieved, there was no difference immediately after the pronoun. But, later, when gender and number information was accessed, coreference was impeded in SPEC sentences as reading times were longer for the SPEC than the NP condition later on in the sentence. Kennison (2003) also replicated the results of “her” with “his”.

Kennison’s (2003) findings contradict Nicol and Swinney (1989) and Clifton *et al.* (1997) as structurally unavailable antecedents were considered as potential subjects since the type of subject influenced reading times. Her findings also contradict Badecker and Straub (2002), as number features appeared to help compose the initial candidate set, while gender mismatch only influenced processing at a later phase. It seemed that the antecedent search ended more quickly when the unavailable candidate differed in number with the pronoun whereas the antecedent search was longer when the subject of the sentence in NP matched the pronoun in gender.

In another experiment, Kennison (2003) aimed to determine whether subject type would influence processing when the discourse context contained an available antecedent for the pronoun as in (28).

- (28) Billy complained about having a stomachache.
- a) Laura watched **him** closely throughout the day.
 - b) Michael watched **him** closely throughout the day.
 - c) They watched **him** closely throughout the day.

The results suggested that when a single highly salient and structurally available antecedent was in discourse context, structurally unavailable antecedents did not influence coreference, which means that when there is a good fit between the antecedent and the pronoun, the process of searching for an antecedent terminates. It appeared that, on the other hand, when no antecedent is available or when there is not a strong fit between the structurally available antecedent and the pronoun, the process of searching for an antecedent continues, and structurally unavailable antecedents can be considered.

Parker (2014) studied how the parser targets specific information in memory, and how that information is extracted to elaborate the sentence representation. The author studied attraction effects in anaphora resolution manipulating gender, number, and animacy. The results for 1-feature mismatch only showed a late slow down in reading times for ungrammatical sentences, and no attraction effects were found. However, for 2-feature mismatch conditions, early and late reading times were facilitated for ungrammatical sentences with attractors when compared to ungrammatical sentences without attractors. Parker (2014) explains that attraction effects are likely to be a consequence of quantitative similarity. Qualitative factors are also important since structural cues are weighted more strongly in retrieval than morphological cues.

Patil *et al.* (2016) thought that reflexive binding may be a very informative phenomenon in understanding the role that grammatical and non-grammatical constraints play in memory. The structural constraints of reflexive binding are relatively clear, and this construction admits manipulations of agreement, distance, and distracting antecedent candidates. They created a model running 1000 simulations of each condition of Sturt's (2003) conditions. Just like Sturt (2003), they found that: retrieval errors on mismatch conditions were higher than in match conditions (mismatch effect), the retrieval errors for both interference conditions, mismatch and match, were higher than for the other 2 conditions (match interference effect), and the retrieval times for both mismatch conditions are longer than the other two match conditions (mismatch effect). On the other hand, they also found results that were not consistent with Sturt (2003): retrieval times for the match interference condition were shorter than for the match condition and shorter than for the mismatch conditions (mismatch interference effect). Patil *et al.* (2016) suggested that the unacceptable candidates in Sturt (2003) could not be good attractors as semantic matching cues are not able to cause attraction if no grammatical cue is involved. In addition, since they were less recently created in representation, they could not have enough strength in memory to be retrieved due to decay factors.

Patil *et al.* (2016) also conducted an eye-tracking experiment. To increase the strength of the inaccessible subject, they used an object pronoun within a relative clause where the inaccessible antecedents were the subject of the clause. Patil *et al.* (2016) found a significant main effect of interference in First Pass and in First Pass Regression Probability. There was also a main effect of match for Rereading times

and Total Reading Times. Thus their results are consistent with Badecker and Straub (2002), but inconsistent with Sturt (2003), Nicol and Swinney (1989), Xiang *et al.* (2009), and Dillon *et al.* (2013). Patil *et al.* (2016) concluded that non-structural cues are crucial for antecedent retrieval so that agreement features such as gender must be included in the set of retrieval cues. Moreover, it seems that strict syntactic constraints on antecedent retrieval are inconsistent with their results, as their results challenged the idea that the parser is infallible for reflexive binding.

5. The present study

The experiment that is reported here is an eye-tracking study, and its main purpose is to investigate how and when the structural constraints of Principle B and agreement cues influence the way nominal antecedents are retrieved from memory.

In this experiment, participants had their eye movements recorded while they read text on a computer screen. Using appropriate software, the researcher can measure the duration of eye fixations (among other measures). This technique is one of the most efficient means linguists have to study language processing. Moreover, it has advantages over the self-paced reading technique because the text can be presented more naturally to the readers (i.e, without segmentation and button pressing).

According to Just and Carpenter (1980), the duration of eye fixations during sentence processing depends on information complexity, that is, the more complex information processing is, the longer the fixation duration in the area where that information is located. These authors make two assumptions: the first is called the *Immediacy Assumption*, which claims that language processing is immediate, that is, a word is processed at the first time it is encountered; the second is called the *Eye-Mind Assumption*, which means that the eye remains fixated on a word as long as the word is being processed. The first assumption is still considered true; however, the second assumption is no longer thought to be true, since a word can still be processed when the eyes are fixated on the next word, which is called the *spillover effect*.

We assume that since overt and redundant agreement marking is often available in languages with rich morphology such as Brazilian Portuguese, speakers will tend to strongly rely on agreement morphology in order to resolve coreference. In congruence with Badecker and Straub

(2002), the hypothesis is that candidates that feature-match the pronouns would initially influence coreference processing, even though they violate Principle B, and that the structural constraints of Principle B would only work later on in binding processing to help the parser select the most adequate antecedent. Therefore, find main effects of structurally unacceptable antecedents at early eye measures (First Fixation Duration) and main effects of structurally acceptable antecedents as late eye measures (Total Fixation Duration) are expected.

In addition, it is hypothesized that memory is sensitive to different types of features. In other words, marked features in the language will be more salient in memory, facilitating memory retrieval. This way, it is expected that due to their markedness in Brazilian Portuguese, feminine and plural features on structurally unacceptable candidates will cause facilitation effects when compared to masculine and singular. It should be mentioned that the markedness of these features is not inherent to them. Plural, for example, is not marked because of its morphology (morpheme –s) or notional plurality, but because of its grammatical number (STAUB, 2009). Plural or feminine is marked in opposition to singular and masculine respectively because the former ones, and not the latter ones, are the *default* features, which are automatic, frequent, and dominant.

Moreover, we hypothesize that as memory decays, sentences in which the structurally acceptable antecedent is linearly distant from the pronoun would have stronger facilitation effects caused by structurally unacceptable antecedents, as they might be more easily retrieved as the antecedents by memory due to recentness (cf. among others, SCHWEPPE, 2013; CHOW *et al.*, 2014).

Participants

Twenty-nine native speakers of Brazilian Portuguese with normal or corrected-to-normal vision participated as volunteers in the experiment. They were undergraduate students of the Federal University of Rio de Janeiro (UFRJ) and were randomly invited to participate in the study, and, as compensation for their work, they receiving three hours of *Cultural-Scientific Activities (Atividades-Científico-Culturais Discentes, AACC)*, which is mandatory for their graduation. All participants were naive with respect to the object of study of the experiment and signed a consent form which stated that the task they would perform would not have any risks to their health and that the results would be eventually

published. Of the twenty-nine participants, five were excluded from analysis as they had less than 80% of their eyes movements recorded. Therefore, the experiment was analyzed using data from the remaining twenty-four participants sixteen female and eight male, with a mean age of 22.6 years (ranging from 18 to 30 years).

Design and materials

There were two independent variables in the experiment. The first one was (i) *structurally acceptable antecedent matching*. In this variable the structurally acceptable antecedent could feature-match/mismatch the pronoun in number. The second one was (ii) *structurally unacceptable antecedent matching*, and the structurally unacceptable antecedent could feature-match/mismatch the pronoun in number.

Besides the independent variables, there were three controls in the experiment: i) the *number of the structurally unacceptable antecedent*, half of the sentences contained plural structurally unacceptable antecedents and the other half singular; ii) the *gender of the structurally unacceptable antecedent*, half of the sentences contained feminine structurally unacceptable antecedents and the other half masculine; iii) the *linear distance between the structurally acceptable antecedent and the pronoun*, half of the sentences contained long linear distance and the other half short. Although the controls could not be considered independent variables, they were taken into account in the analysis of the experiment.

The experiment had two on-line dependent variables: (i) the First Fixation Duration and (ii) the Total Fixation Duration at the pronoun areas.

Each of the four lists, which were created using a Latin Square, was pseudo-randomized and contained sixteen experimental sentences and thirty-two fillers. Four sentences from each experimental condition were in each list. Each sentence of the experiment was accompanied by an off-line yes-or-no comprehension question. The filler questions were balanced between yes and no answers, while all the experimental sentences had yes answers.

Each experimental trial contained a structurally acceptable antecedent (masculine/feminine, singular/plural) in the main clause, followed by a structurally unacceptable antecedent, which was the subject of a relative clause, followed by a 3rd person pronoun (“ele/ela/ eles/elas”), which were the direct objects of the relative clauses. One can find an example sentence below:

(29) Short distance between the structurally acceptable antecedent and the pronoun:

TABLE 1 – Sample of the experimental materials used for short distance conditions

	Feminine structurally unacceptable antecedent	Masculine structurally unacceptable antecedent
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent match	As <u>professoras</u> contraram a <u>faxineira</u> que xinga ela com muitos palavrões. (“The[fem, pl] <u>teachers</u> [fem,pl] hired[pl] the[fem,sg] <u>cleaner</u> [fem,sg] who curse[sg] her with a lot of bad words.”)	O <u>médico</u> elogiou os <u>enfermeiros</u> que chamam eles quando necessário. (“The[masc,sg] <u>doctor</u> [masc, sg] complimented[sg] the[masc, pl] <u>nurses</u> [masc, pl] who called them when necessary.”)
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent mismatch	As <u>professoras</u> contraram as <u>faxineiras</u> que xingam ela com muitos palavrões. (“The[fem, pl] <u>teachers</u> [fem,pl] hired[pl] the[fem,pl] <u>cleaners</u> [fem,pl] who curse[sg, pl] her with a lot of bad words.”)	O <u>médico</u> elogiou o <u>enfermeiro</u> que chama eles quando necessário. (“The[masc,sg] <u>doctor</u> [masc, sg] complimented[sg] the[masc, sg] <u>nurse</u> [masc, sg] who called them when necessary.”)
Structurally acceptable antecedent match/ structurally unacceptable antecedent match	As <u>professoras</u> contraram as <u>faxineiras</u> que xingam elas com muitos palavrões. (“The[fem, pl] <u>teachers</u> [fem,pl] hired[pl] the[fem,pl] <u>cleaners</u> [fem,pl] who curse[sg, pl] them with a lot of bad words.”)	O <u>médico</u> elogiou o <u>enfermeiro</u> que chama ele quando necessário. (“The[masc,sg] <u>doctor</u> [masc, sg] complimented[sg] the[masc, sg] <u>nurse</u> [masc, sg] who called him when necessary.”)
Structurally acceptable antecedent match/ structurally unacceptable antecedent mismatch	As professoras contraram a faxineira que xinga elas com muitos palavrões. (“The[fem, pl] <u>teachers</u> [fem,pl] hired[pl] the[fem,sg] <u>cleaner</u> [fem,pl] who curse[sg, sg] them with a lot of bad words.”)	O <u>médico</u> elogiou os <u>enfermeiros</u> que chamam ele quando necessário. (“The[masc,sg] <u>doctor</u> [masc, sg] complimented[sg] the[masc, pl] <u>nurses</u> [masc, pl] who called him when necessary.”)

(30) Short distance between the structurally acceptable antecedent and the pronoun:

TABLE 2 – Sample of the experimental materials used for long distance conditions

	Feminine structurally unacceptable antecedent	Masculine structurally unacceptable antecedent
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent match	<p>As <u>cozinheiras</u> viram vários problemas graves e razões sérias para despedir a <u>lavadeira</u> que olha ela com antipatia. (“The[fem, pl] <u>cooks</u>[fem, pl] saw several serious problems and reasons to fire the[fem, sg] <u>washerwoman</u>[fem, sg] who look[sg] at her with antipathy.”)</p>	<p>Os <u>mecânicos</u> viram bem poucas qualificações e vantagens para contratar o <u>borracheiro</u> que chama ele na oficina. (“The[masc, pl] <u>mechanics</u> [masc, pl] saw a few qualifications and advantages to hire the[masc, sg] <u>tire repairman</u>[masc, sg] who call[sg] him in the repair shop.”)</p>
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent mismatch	<p>As <u>cozinheiras</u> viram vários problemas graves e razões sérias para despedir as <u>lavadeiras</u> que olham ela com antipatia. (“The[fem, pl] <u>cooks</u>[fem, pl] saw several serious problems and reasons to fire the[fem, pl] <u>washerwomen</u>[fem, pl] who look[pl] at her with antipathy.”)</p>	<p>Os <u>mecânicos</u> viram bem poucas qualificações e vantagens para contratar os <u>borracheiros</u> que chamam ele na oficina. (“The[masc, pl] <u>mechanics</u> [masc, pl] saw a few qualifications and advantages to hire the[masc, pl] <u>tire repairmen</u>[masc, pl] who call[pl] him in the repair shop.”)</p>
Structurally acceptable antecedent match/ structurally unacceptable antecedent match	<p>As <u>cozinheiras</u> viram vários problemas graves e razões sérias para despedir as <u>lavadeiras</u> que olha elas com antipatia. (“The[fem, pl] <u>cooks</u>[fem, pl] saw several serious problems and reasons to fire the[fem, pl] <u>washerwomen</u>[fem, pl] who look[pl] at them with antipathy.”)</p>	<p>Os <u>mecânicos</u> viram bem poucas qualificações e vantagens para contratar os <u>borracheiros</u> que chamam eles na oficina. (“The[masc, pl] <u>mechanics</u> [masc, pl] saw a few qualifications and advantages to hire the[masc, sg] <u>tire repairmen</u>[masc, pl] who call[pl] them in the repair shop.”)</p>
Structurally acceptable antecedent match/ structurally unacceptable antecedent mismatch	<p>As <u>cozinheiras</u> viram vários problemas graves e razões sérias para despedir a <u>lavadeira</u> que olha elas com antipatia. (“The[fem, pl] <u>cooks</u>[fem, pl] saw several serious problems and reasons to fire the[fem, sg] <u>washerwoman</u>[fem, sg] who look[sg] at them with antipathy.”)</p>	<p>Os <u>mecânicos</u> viram bem poucas qualificações e vantagens para contratar o <u>borracheiro</u> que chama ele na oficina. (“The[masc, pl] <u>mechanics</u> [masc, pl] saw a few qualifications and advantages to hire the[masc, sg] <u>tire repairman</u>[masc, sg] who call[sg] him in the repair shop.”)</p>

Procedure

The experiment was conducted at the laboratory of experimental research (LAPEX) at the Federal University of Rio de Janeiro (UFRJ) in Rio de Janeiro, Brazil. The eye-tracking software used in this experiment was *Tobii Studio™ TX 300*, which requires an initial individual calibration at the beginning of the procedure for the eye-tracker to be able to monitor the participant's pupils during the reading task. The participants were instructed to sit comfortably and were given written and oral task instructions. After that, the calibration process would start, followed by a short practice session with filler sentences so that the experimenter could check whether the participants understood the task and were performing it at a natural speed. Finally, the experimenter left the participants alone in a quiet room without distraction. Each sentence of the experiment would appear in whole on the computer screen. The participants could read each sentence however many times that was necessary; however they were instructed to read each sentence as fast as they could while also paying attention to meaning. After reading a sentence, the participants would press the space bar to continue to a comprehension question about the sentence that was just read. Subjects answered by fixating their eyes on one of the options, "Yes" or "No". Each participant randomly performed one of the four lists of the experiment. The duration of the experiment was of approximately twenty minutes.

Analysis

The reading time data were extracted using *Tobii Fixation Filter*, which is the default fixation algorithm in *Tobii Studio™ 2.X version 2.2*. Approximately 19% of the data were lost due to calibration issues. Therefore, due to the small sample of the test, we decided not to perform any outlier trimming. Our raw data came with a positively skewed non-normally distributed population (*Shapiro Test*: $W=0.893$, $p<0.05$ for First Fixation; and $W= 0.619$, $p<0.05$ for Total Fixation). We believe that this was a consequence of the small sample and the missing data. We did not transform our data to achieve normality, because we decided to analyze the experiment with a linear mixed-effect model (LMM), which is a statistical model that does not use mean data, as it does not average across individual responses and it can also cope with unbalanced data (LO; ANDREWS, 2015). Moreover, we were concerned that data

transformation could bias the results. All dependent variables were *within-subjects* and the statistical analysis was carried on R⁶ software, using *plotrix*,⁷ *lmer Test*,⁸ and *gplots*⁹ packages.

Although there were only two independent variables as mentioned before, we also included the three controlled variables in our analysis in order to reduce the error residual in our LMM statistical model. Therefore, five variables were analyzed: a) *structurally acceptable antecedent matching* (matching or mismatching), b) *structurally unacceptable antecedent matching* (matching or mismatching), c) *number of the structurally unacceptable antecedent* (singular or plural), d) *gender of the structurally unacceptable antecedent* (feminine or masculine), e) *distance between the structurally acceptable antecedent and the pronoun* (short or long).

Results

Means as well as standard errors of First Fixation Duration and Total Fixation Duration at the pronoun area are reported for each condition in Tables 1 and 2:

- (31) Short distance between the structurally acceptable antecedent and the pronoun:

⁶ R CORE TEAM. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria, 2016. (<https://www.r-project.org/>)

⁷ LEMON, J. Plotrix: a package in the red light district of R. *R-News*, R Foundation, v. 6, n. 4, p. 8-12, 2006.

⁸ Alexandra Kuznetsova, Per Bruun Brockhoff and Rune Haubo Bojesen Christensen (2015). *lmerTest*: Tests in Linear Mixed Effects Models. R package version 2.0-29 (<https://CRAN.R-project.org/package=lmerTest>)

⁹ Gregory R. Warnes, Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc Schwartz and Bill Venables (2015). *gplots*: Various R Programming Tools for Plotting Data. R package version 2.16.0. (<https://CRAN.R-project.org/package=gplots>)

TABLE 3 – First Fixation Duration means and standard errors in milliseconds for short distance experimental conditions

	Structurally unacceptable number	Eye measure	Feminine structurally unacceptable antecedent	Masculine structurally unacceptable antecedent
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent match	Singular	First Fixation	297 (107)	266 (148)
		Total Fixation	375 (128)	542 (253)
	Plural	First Fixation	350 (195)	253 (76)
		Total Fixation	547 (264)	462 (294)
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent mismatch	Singular	First Fixation	294 (179)	219 (67)
		Total Fixation	379 (187)	499 (301)
	Plural	First Fixation	285 (171)	349 (161)
		Total Fixation	323 (250)	531 (251)
Structurally acceptable antecedent match/ structurally unacceptable antecedent match	Singular	First Fixation	298 (105)	251 (123)
		Total Fixation	431 (172)	478 (207)
	Plural	First Fixation	259 (40)	279 (53)
		Total Fixation	270 (60)	395 (189)
Structurally acceptable antecedent match/ structurally unacceptable antecedent mismatch	Singular	First Fixation	256 (77)	263 (76)
		Total Fixation	421 (257)	599 (374)
	Plural	First Fixation	281 (48)	327 (118)
		Total Fixation	630 (296)	317 (130)

(32) Long distance between the structurally acceptable antecedent and the pronoun:

TABLE 4 – First Fixation Duration means and standard errors in milliseconds for long distance experimental conditions

	Structurally unacceptable number	Eye measure	Feminine structurally unacceptable antecedent	Masculine structurally unacceptable antecedent
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent match	Singular	First Fixation	344 (121)	316 (142)
		Total Fixation	585 (322)	616 (360)
	Plural	First Fixation	242 (81)	277 (69)
		Total Fixation	522 (148)	478 (202)
Structurally acceptable antecedent mismatch/ structurally unacceptable antecedent mismatch	Singular	First Fixation	326 (121)	294 (126)
		Total Fixation	442 (166)	897 (1425)
	Plural	First Fixation	329 (180)	379 (217)
		Total Fixation	614 (515)	732 (617)
Structurally acceptable antecedent match/ structurally unacceptable antecedent match	Singular	First Fixation	339 (240)	281 (132)
		Total Fixation	576 (467)	328 (154)
	Plural	First Fixation	263 (52)	258 (44)
		Total Fixation	382 (111)	349 (133)
Structurally acceptable antecedent match/ structurally unacceptable antecedent mismatch	Singular	First Fixation	288 (158)	284 (79)
		Total Fixation	489 (266)	454 (178)
	Plural	First Fixation	301 (144)	305 (87)
		Total Fixation	381 (145)	555 (312)

LMM was created with the help of *lmerTest* package with the following Fixed Effects: i) *structurally acceptable antecedent matching*; ii) *structurally unacceptable antecedent matching*; iii) *number of the structurally unacceptable antecedent*; iv) *gender of the structurally unacceptable antecedent*; v) *distance between the structurally acceptable antecedent and the pronoun*. On the other hand, the Random Effects were: i) *participants*; and ii) *items*.

For First Fixation Duration, using the *anova* () function in our model, it was found a significant effect of the interaction between the variables *structurally unacceptable antecedent* and *structurally unacceptable antecedent number*: $F(1, 258)=6.248, p=0.013^*$. In addition, there was an important trend towards statistical significance in the interaction between *structurally unacceptable antecedent gender* and *number*: $F(1, 265)=3.44, p=0.064$; and a moderately significant interaction between *structurally unacceptable antecedent number* and *distance between the structurally acceptable antecedent and the pronoun*: $F(1, 264)=2.46, p=0.117$.

For Total Fixation Duration, a linear mixed-effect model was also created with the help of *lmerTest* package. Its fixed and random effects were the same of the First Fixation Duration model. By using the *anova* () function, we found a significant main effect of structurally acceptable antecedent matching: $F(1, 254) = 4.046, p=0.045^*$ and slight trend towards significance in the interaction between *structurally acceptable antecedent* and *linear distance*: $F(1, 253)=3.556, p=0.060$.

In order to figure out which pairs of conditions were significantly different, bar plots with *Tukey Tests Results for 95% confidence intervals (CI)* were also created with the help of *ggplot2* package.

Structurally acceptable antecedents that feature-matched the pronouns had faster reading times when compared to structurally acceptable antecedents that mismatched the pronouns for First Fixation Duration, although this effect was probably not statistically significant ($\beta=-21, CI [-49, 6], p=0.140$), and for Total Fixation Duration, with a statistically significance difference ($\beta=-91, CI [-176, 5], p=0.036^*$), as one can see in Figures 2 and 3 respectively.

FIGURE 2 – 95% CI barplot for First Fixation Duration of structurally acceptable antecedent matching

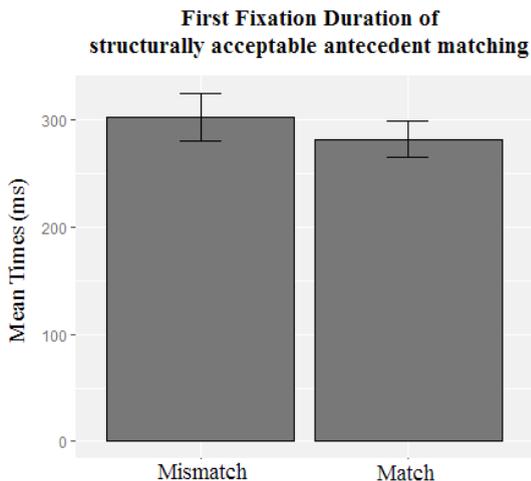
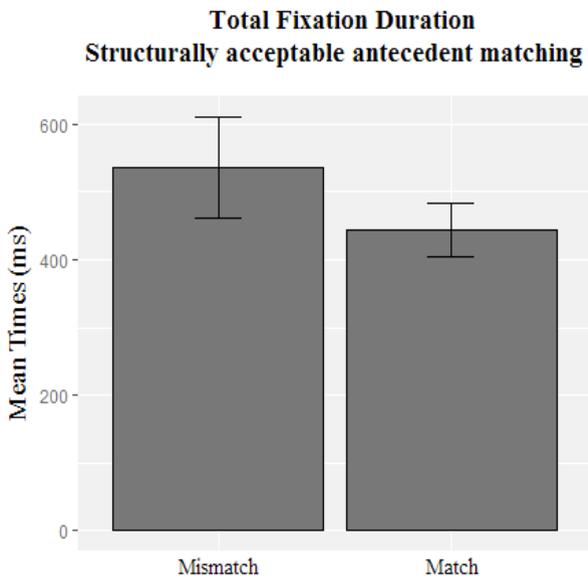


FIGURE 3 – 95% CI barplot for Total Fixation Duration of structurally acceptable antecedent matching



Even though this was only a near-marginal statistical significance, Figure 4 illustrates that First Fixation reading times at the pronoun area were slower when there was a long distance between the pronoun and the structurally acceptable antecedent when compared to short distance, ($\beta=18$, CI [-9, 46], $p=0.187$). Figure 5 shows an established trend towards significance for Total Fixation Duration: ($\beta=70$, CI [-15, 155], $p=0.106$).

FIGURE 4 – 95% CI barplot for First Fixation Duration of the linear distance between the structurally acceptable antecedents and the pronouns

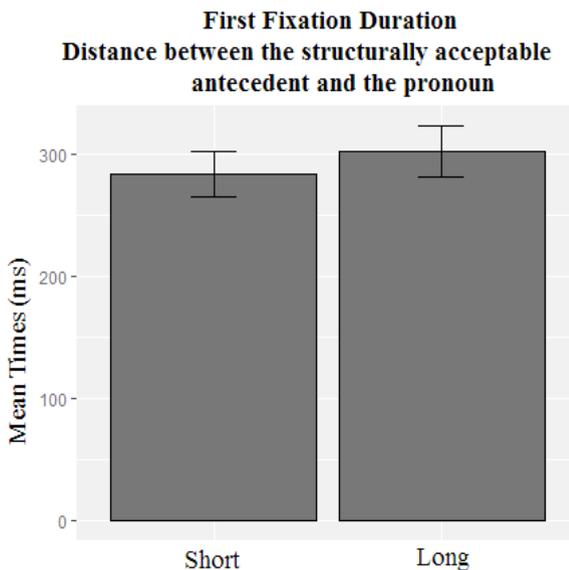
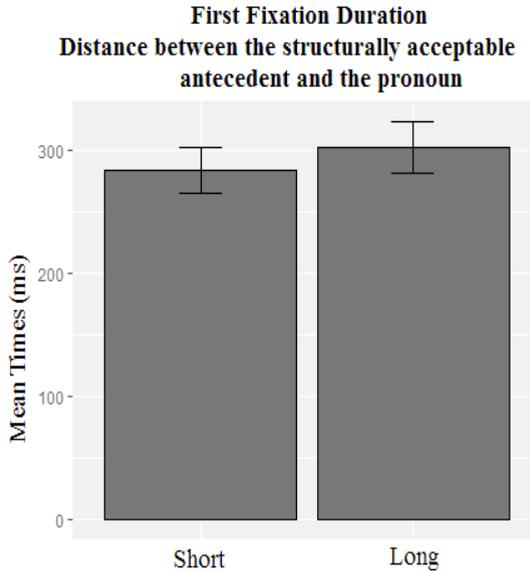


FIGURE 5 – 95% CI barplot for Total Fixation Duration of the linear distance between the structurally acceptable antecedents and the pronouns



For First Fixation Duration, one can see in Figure 6, there was near-marginal significance: a) structurally unacceptable antecedents in the plural that feature-mismatched the pronoun had longer reading times when compared to structurally unacceptable antecedents that matched the pronouns ($\beta=41$, CI [-11, 94], $p=0.180$); and an apparent trend towards significance: b) singular structurally unacceptable antecedents that mismatched the pronouns had faster First Fixation times than plural structurally unacceptable antecedents ($\beta=-38$, CI [-90, 13], $p=0.215$).

FIGURE 6 – 95% CI barplot for First Fixation Duration of structurally unacceptable antecedents and number

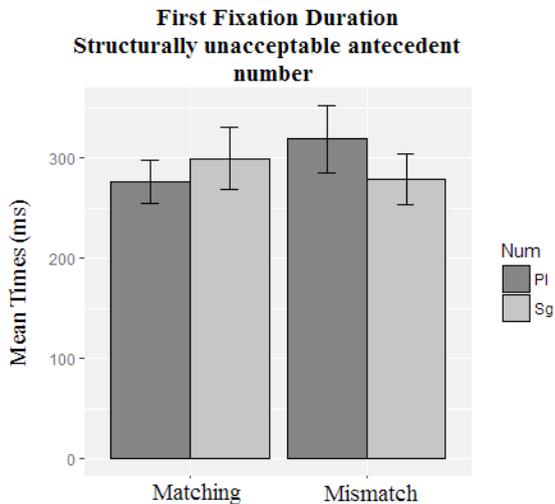
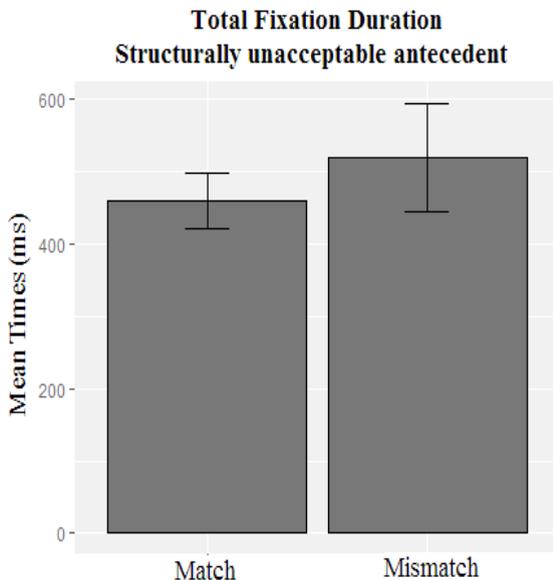


FIGURE 7 – 95% CI barplot for Total Fixation Duration of structurally unacceptable antecedents



Similarly, for Total Fixation Duration, Figure 7 illustrates near-marginal significance for longer reading times when the structurally unacceptable antecedent mismatches the pronoun than when it matches: ($\beta=57$, CI [-27, 143], $p=0.186$).

Discussion

This research aimed investigated how pronouns retrieve their antecedents in Brazilian Portuguese, a morphologically rich language. It was hypothesized that due to the redundant overt agreement marking in Brazilian Portuguese, readers would strongly rely on agreement cues in order to retrieve pronominal antecedents. Therefore, we expected that structurally unacceptable antecedents would be initially considered as potential antecedents, even they violate the structural constraints of Principle B, which would only influence coreference processing later, helping the parser to select the most adequate antecedent from the initial candidate set. The results of the present experiment seem to corroborate that hypothesis, providing evidence in favor of Badecker and Straub (2002). The results of the LMM for First Fixation Duration, which measures early processing, showed effects of structurally unacceptable antecedents, and agreement features such as number and gender, as well as the linear distance between the structurally acceptable antecedents and the pronouns. However, it should be mentioned that 95% confidence intervals for our conditions in First Fixation Duration also showed effects of structurally acceptable antecedents, that is, effects of Principle B. On the other hand, our LMM for Total Fixation only showed effects of structurally acceptable antecedents and linear distance, although the 95% confidence intervals for our conditions in Total Fixation Duration also showed effects of structurally unacceptable antecedents.

It is important to note the 95% confidence intervals showed that structurally unacceptable antecedents that feature-matched the pronouns were responsible for faster coreference processing in both First Fixation Duration and Total Fixation Duration, which might be evidence that these candidates are actually being initially retrieved as antecedents. According to Dillon (2013), facilitatory effects of structurally unacceptable antecedents might be evidence that antecedents are retrieved through a content-addressable memory. In other words, structurally unacceptable antecedents can cause interference effects in memory due to the fact that they partially match the content cues of the pronouns, leading to

erroneous retrieval of them as antecedents, which in CAM is known as similarity-based interference.

Therefore, these results evidence contradicting the hypothesis that structural constraints work as an initial filter in binding processing, blocking the influence of structurally unacceptable candidates, as claimed by Nicol and Swinney (1989), Clifton *et al.* (1997), Sturt (2003), Leitão (2008), Xiang *et al.* (2009), Oliveira *et al.* (2012), Dillon *et al.* (2013) and Chow *et al.* (2014). These results also contradict Kennison (2003), as it seems that gender agreement is also important at initial stages of coreference processing.

Curiously, at early processing, it was also found that even when the structurally unacceptable antecedents mismatched the pronouns, coreference processing is affected. In this case, structurally unacceptable antecedents in the singular tend to facilitate coreference when compared to the ones in the plural. This result apparently contradicts CAM, as this model posits that only partial matches can cause similarity-based interference effects, and since structurally unacceptable antecedents that mismatch the pronouns do not have any content cue matching the pronoun, it should not be taken into account by memory as a potential antecedent.

It is noteworthy that it was not found strong evidence to support the second hypothesis, that memory is sensitive to different types of agreement features. It was expected that feminine structurally acceptable antecedents would influence coreference more than the masculine ones due to the fact that feminine features are marked in Brazilian Portuguese. The same applies for number, when comparing plural to singular. Our results did not show any difference between masculine and feminine, but it suggests that, as already mentioned above, structurally unacceptable antecedents in the singular that mismatched the pronouns were responsible for shorter reading times at the pronoun area than plural features. Interestingly, plural features did not cause facilitatory effects as we expected, but rather increased reading times. However, one must remember that in that condition, the structurally unacceptable antecedents mismatched the pronouns; therefore, for CAM, they could not be considered as potential antecedents. Essentially, the structurally unacceptable antecedents only facilitate coreference when they feature-match the pronouns. And in the relevant condition, this was not the case. Nonetheless, we should not assume that singular and plural features

did not behave the same in this condition. It is likely that structurally unacceptable antecedents in the plural that mismatched pronouns can bring difficulties to the processor due to the salience of plural features in memory. It might be the case that the parser knows that this specific candidate could not be the adequate antecedent, but at the same time the salience of plural disturbs memory, even though those features do not match the pronoun.

As expected, another piece of evidence in favor of CAM is the decay effects we found in both First Fixation Duration and Total Fixation Duration, that is, short linear distance between the structurally acceptable antecedents and the pronouns facilitated coreference when compared to long distance. Again, the hypothesized reason for this is that the more recent an item was stored in memory; the easier it is to retrieve it (cf. among others, Schweppe, 2013; Chow et al, 2014).

Finally, it was found that structurally acceptable antecedents that matched the pronouns facilitated coreference in both First Fixation Duration and Total Fixation Duration. Once structurally acceptable candidates totally match the content cues of the pronouns (both agreement and structural ones), they can clearly be considered as the best antecedents. Nevertheless, it was not expected to find this result in First Fixation Duration, as it was hypothesized that the structural constraints of Principle B would only work at late processing phases, as in Total Fixation Duration. However, as this effect did not appear in our LMM for First Fixation as a significant effect or a trend towards significance, we can continue to support our hypothesis.

6 Conclusion

This research filled a gap in the literature on coreference processing by showing how agreement cues influence antecedent retrieval in a language with rich morphology and how memory seems to be sensitive to different types of features. In summary, the results suggest that agreement cues, structural constraints, and decay effects can influence coreference from early to late processing stages. However, it seems that agreement cues play a major role at early pronominal coreference processing, while structural constraints play a major role at later processing. Decay appears to influence processing regardless of particular agreement feature or structural constraint. Moreover, structurally unacceptable antecedents that feature match the pronoun in

gender and number facilitated coreference processing. It was suggested that these structurally unacceptable antecedents might be erroneously retrieved from memory as potential antecedents as a result of similarity-based interference.

The results reported here provide evidence in favor of CAM as it was found decay and similarity-interference effects in antecedent retrieval from memory; however, it seems that this model needs some adjustments in order to explain how candidates that mismatched the pronouns could interfere in coreference processing, how singular and plural features can be distinguished in memory, and what the consequences are for these effects in sentence processing. Thus future studies that seek to compare different types agreement cues in antecedent memory retrieval can be helpful in order to better understand how memory retrieval and language processing are integrated.

Acknowledgements

I am grateful for the reviewers' comments and suggestions, which enormously contributed to this paper. I would like to thank William Matchin for his helpful discussions on the results and his generous proofreading. I am also grateful for the participants that were involved in the experiment reported here. Finally, it should be mentioned that without the financial aid given by *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq), this research would not be possible.

References

BADECKER, W.; STRAUB, K. The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Washington, v. 28, p.748-769, 2002. DOI: <http://dx.doi.org/10.1037//0278-7393.28.4.748>.

BOCK, K.; MILLER, C. Broken agreement. *Cognition Psychology*, Elsevier, v. 23, p. 45-93, 1991.

CHOMSKY, N.; LASNIK, H. *The theory of principles and parameters*. In: JACOBS, J. et alii (Org.). *Syntax, an International Handbook of Contemporary Research*. Berlin; New York: Walter de Gruyter, 1993. DOI: <http://dx.doi.org/10.3389/fpsyg.2014.00630>.

CHOW, W.; SHEVAUN, L.; PHILLIPS, C. Immediate sensitivity to structural constraints in pronoun resolution. *Frontiers in Psychology*, Frontiers Media, v. 5, article 630, 2014. <http://doi.org/10.3389/fpsyg.2014.00630>.

CLIFTON, C.; FRAZIER, L.; DEEVY, P. Feature manipulation in sentence comprehension. *Rivista di Linguistica, Laboratorio di Linguistica 'Giovanni Nencioni'*, v. 11, p.11-39, 1997.

DILLON, B.; MISHLER, A.; SLOGGETT, S.; PHILLIPS, C. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, Elsevier, v. 69, p.85-103, 2013. DOI: <http://dx.doi.org/10.1016/j.jml.2013.04.003>.

JUST, M. C.; CARPENTER, P. A. A theory of reading: from eye fixations to comprehension. *Psychological Review*, Elsevier, v. 87, n. 4, p. 329-354, 1980.

KENNISON, S. M. Comprehending the pronouns *her*, *him*, and *his*: implications for theories of referential processing. *Journal of Memory and Language*, Elsevier, v. 49, p.335-352, 2003. DOI: [http://dx.doi.org/10.1016/S0749-596X\(03\)00071-8](http://dx.doi.org/10.1016/S0749-596X(03)00071-8).

LAGO, S. Agreement processes in Spanish comprehension. *Journal of Memory and Language*, v. 82, p.133-149, 2015. DOI: <http://dx.doi.org/10.1016/j.jml.2015.02.002>

LEITÃO, M.; PEIXOTO, P.; SANTOS, S. Processamento da correferencial intrasentencial em PB. *Veredas Online*, Juiz de Fora, UFJF, v. 2, p.50-61, 2008.

LEWIS, R. L.; VASISHTH, S. An activation-based Model of sentence processing as skilled memory retrieval. *Cognitive Science*, Wiley Online Library, v. 29, p.375-419, 2005.

LEWIS, R. L.; VASISHTH, S.; VAN DYKE, J. A. Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences (Regul.Ed)*, Elsevier, v. 10, p.44-54, 2006. DOI: <http://dx.doi.org/10.1016/j.tics.2006.08.007>.

LO, S.; ANDREWS, S. To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Frontiers of Psychology*, Frontiers Media, v. 6, 2015. DOI: <http://dx.doi.org/10.3389/fpsyg.2015.01171>.

McELREE, B. Sentence comprehension is mediated by content addressable memory structures. *Journal of Psycholinguistic Research*, Springer Link, v. 29, p.111-123, 2000.

McELREE, B.; FORAKER, S.; DYER, L. Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, Elsevier, v. 48, p. 67-91, 2003.

NICOL, J.; SWINNEY, D. A. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, Springer Link, v. 18, p.5-19, 1989.

OLIVEIRA, R.; LEITÃO, M.; HENRIQUE, J. A influência dos antecedentes vinculados e não vinculados no processamento da anaphora “a si mesmo”. *Revista Linguística*, Rio de Janeiro, UFRJ, v. 8, n. 2, p. 71-85, 2012. DOI: <http://dx.doi.org/10.17074/linguistica.v8i2.586>.

PARKER, D. *The cognitive basis for encoding and navigating linguistic structure*. 2014. Dissertation (PhD) – University of Maryland, 2014.

PATIL, U.; VASISHTH, S.; LEWIS, R. Retrieval interference in syntactic processing: The case of reflexive binding in English. *Frontiers of Psychology*, Frontiers Media, v. 7, Artigo 329, 2016. DOI: <http://dx.doi.org/10.3389/fpsyg.2016.00329>.

PHILLIPS, C; WAGERS, M; LAU, E. Grammatical illusions and selective fallibility in real-time language comprehension. In: RUNNER, Jeffrey (Ed.) *Experiments at the Interfaces*. Emerald Group Publishing Limited, Wagon Lane, UK, 2011. (Série Syntax and Semantics, v. 37, p.147-180).

SCHWEPPE, J. Distance Effects in Number Agreement. *Discourses Processes*, Taylor & Francis Online, v. 50, n. 8, p.531-556, 2013. DOI: <http://dx.doi.org/10.1080/0163853X.2013.841074>.

STAUB, A. On the interpretation of the number attraction effect: Response time evidence. *Journal of Memory and Language*, Elsevier, v. 60, p. 308-327, 2009. DOI: <http://dx.doi.org/10.1016/j.jml.2008.11.002>.

STURT, P. The time-course of the application of binding constraints in reference. *Journal of Memory and Language*, v.48, p.542-562, 2003. DOI: [http://dx.doi.org/10.1016/S0749-596X\(02\)00536-3](http://dx.doi.org/10.1016/S0749-596X(02)00536-3).

VAN DYKE, J.; Mc ELREE, B. Retrieval interference in sentence comprehension. *Journal of Memory and Language*, Elsevier, v. 55, p.157-166, 2006. DOI: <http://dx.doi.org/10.1016/j.jml.2006.03.007> .

WAGERS, M.; LAU, E.; PHILLIPS, C. Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, Elsevier, v. 61, p. 206-237, 2009. DOI: <http://dx.doi.org/10.1016/j.jml.2009.04.002>.

XIANG, M.; DILLON, B.; PHILLIPS, C. Illusory licensing effects across dependency types: ERP evidence. *Brain & Language*, Elsevier, v. 108, p.40-55, 2009. DOI: <http://dx.doi.org/10.1016/j.bandl.2008.10.002>.

A behavioral study to investigate the processing routes of grammatical gender in Brazilian Portuguese

Um estudo comportamental para investigar as rotas do processamento do gênero gramatical do português brasileiro

Natália C. A. de Resende

Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Minas Gerais / Brasil

natcarol2000@gmail.com

Mailce Borges Mota

Universidade Federal de Santa Catarina, Florianópolis, Santa Catarina / Brasil

CNPq

mailce.mota@ufsc.br

Abstract: In this paper, we present a behavioral study aimed at investigating whether inanimate grammatical gender transparent nouns (feminine nouns ending in *-a* and masculine nouns ending in *-o*) and inanimate grammatical gender opaque nouns (other endings) in Brazilian Portuguese are processed by the same or by distinct cognitive mechanisms. To this purpose, 19 subjects participated in a grammatical gender agreement task in two conditions: between a determiner (definite article) and a noun (condition 1), and a noun and an adjective (condition 2). Factors such as frequency of the nouns and adjectives (high vs. low) as well as their phonological form (gender transparent vs. gender opaque) were manipulated. Results show that, in both conditions, frequency is a strong predictor of faster reaction times, suggesting that gender transparent and gender opaque forms can be stored in memory. This pattern holds for both conditions. We interpret this finding as evidence for the single-mechanism view of language processing.

Keywords: grammatical gender; Brazilian Portuguese.

Resumo: Este artigo apresenta um estudo comportamental que teve como objetivo investigar se substantivos inanimados do português do Brasil, transparentes quanto ao seu gênero gramatical (substantivos femininos terminados em *-a* e masculinos terminados em *-o*) e opacos (outras terminações) são ou não processados por mecanismos cognitivos distintos. Para tanto, uma tarefa de concordância de gênero foi executada por 19 sujeitos em duas condições: entre um artigo definido e um substantivo (condição 1) e entre um substantivo e um adjetivo (condição 2). Fatores como a frequência dos substantivos e adjetivos (alta *vs.* baixa) e forma fonológica (transparente *vs.* opaca) foram manipulados. Os resultados mostram que, em ambas as condições, a frequência é um forte preditor de respostas mais rápidas, sugerindo que tanto formas transparentes quanto formas opacas podem ser armazenadas na memória. Este padrão foi encontrado em ambas as condições. Estes resultados foram interpretados como evidência para a visão unitária de processamento da linguagem.

Palavras-chave: gênero gramatical; português do Brasil.

Recebido em: 10 de dezembro de 2016.

Aprovado em: 6 de abril de 2017.

1 Introduction

One of the most intriguing phenomena of the human mind is its ability to learn and process language. Although language processing seems quite simple given the speed and automaticity with which the human mind deals with it, how the brain stores and processes language regularities and irregularities has been largely debated within the Psycholinguistic research field (for example, PINKER, 1991, 1999; ULLMAN, 2001; BYBEE, 1985, 1995; MACWHINNEY, 2001; CHOMSKY, 1981; JACKENDOFF, 2002).

On one view, linguistic processing relies on a dual mechanism (PINKER, 1991, 1999; ULLMAN, 2001): a mental lexicon specialized in the storage of simplex/irregular lexical forms (PINKER, 1991, 1999; ULLMAN, 2001) and a different mechanism specialized in the processing of complex/decomposable lexical forms. On a second view (e. g. BYBEE, 1985, 1995; MCWHINNEY, 2001), simplex as well as complex forms are retrieved from and processed by a single associative memory mechanism.

Both the dual and the single mechanism views discuss the factors that influence the processing of complex and simplex forms. According to dual models of language processing, only simplex/idiosyncratic forms are sensitive to frequency effects. Regular/complex forms, however, usually are not subject to such effect since they are processed by a cognitive system specialized in the computation of rules. The single mechanism view, on the other hand, predicts that both categories of lexical items (simplex and complex forms) are subject to frequency effects as they are all processed and retrieved from an associative lexical memory.

A series of studies (FLEISCHHAUER; CLAHSSEN, 2012; VAN DER LELY; ULLMAN, 2001; PRADO; ULLMAN, 2009, PRASADA; PINKER, 1993; CLAHSSEN, ROTHWEILER; WOEST, 1992; MARCUS *et al.*, 1995) has been proposed to investigate the processing routes of regular vs. irregular inflectional morphology focusing, mainly, on the verbal morphology of languages. For the English language, for instance, many studies (NEWMAN; ULLMAN, 2007; ULLMAN, 2001, 2004, 2005; PRASADA; PINKER, 1993) have shown that the processing of high-frequency and low-frequency regular forms are not correlated with their frequencies, but that speakers of English process high-frequency irregular verbs faster than low-frequency irregular verbs. These findings suggest that speakers of English activate a computational mechanism to process regular forms (by concatenating a base to a suffix: *walk + ed* in language production or decomposing the items into their morphological constituents in language comprehension), but rely on memory to process irregular forms. Research on languages other than English has shown a similar pattern (BOWDEN, 2007; MORGAN-SHORT *et al.*, 2010; CLAHSSEN; EISENBEISS; SONNENSTUHL, 1997). Other findings, however, contradict such claims. Alegre and Gordon (1999), Baayens *et al.* (1997) and Baayen, Tweedie and Schreuder (2002), for instance, have shown that regular forms are also sensitive to surface frequency effects, suggesting that they can also be stored in memory as simplex forms.

Given the debate in the literature, the question that arises is: If regular forms can also be stored, what kinds of structures can be stored and what kinds of structures can be computed? An answer to this question is still far from being trivial. We believe that not only verbal morphology can provide good insights to this debate. In this paper, we test the hypotheses of the dual and single mechanism models by investigating another aspect of morphology: the processing of grammatical gender of

inanimate nouns. The language of our investigation is Portuguese, more specifically Brazilian Portuguese which, so far, has not received much attention on the part of researchers interested in investigating the routes of language processing. The research question we address in the present study is the following: Is grammatical gender agreement in Brazilian Portuguese governed by a dual or by a single cognitive mechanism?

Before presenting a detailed description of the method adopted in the present study, we think it is worth presenting a review of the literature focusing on the cognitive mechanisms involved in the processing of grammatical gender and a brief (but complete) outline of the grammatical gender of Portuguese with the purpose of clarifying how the grammatical gender system of Portuguese can provide subsidies for a study on the balance between storage and computation.

1.1. Literature Review

Grammatical gender, unlike semantic gender, does not refer to the biological sex of the referent, but represents, instead, a formal, intrinsic feature of inanimate nouns (CORBET, 1991). In Portuguese, as in other Romance languages, animate and inanimate nouns are distributed into gender categories: feminine and masculine. For instance, the word *cultura* (“culture”) does not have an animate referent in the world, but it is a feminine noun. Similarly, the word *pensamento* (“thought”) does not have an animate referent in the world, but it is a masculine noun.

In Portuguese, as a general rule, nouns ending in *-a* tend to be feminine, and nouns ending in *-o* tend to be masculine. However, many other nouns with other endings can be of either grammatical gender. Hence, a computational mechanism could be at work to process the grammatical gender rule (nouns ending in *-a* are feminine, nouns ending in *-o* are masculine) and a distinct mechanism, an associative memory, would be responsible for storing and processing the grammatical gender of opaque forms.

Analogous to the debate focusing on the routes of verbal morphology processing, researchers interested in the processing of grammatical gender have also examined whether the grammatical gender of inanimate nouns is represented and processed by a dual or a single cognitive mechanism. A number of studies (e.g. AFONSO *et al.*, 2014; GOLLAN; FROST, 2001; HOHFELD, 2006) in languages other than Portuguese suggest that there is a dual route to process

grammatical gender. A study in Hebrew has offered evidence in favor of such claim. Gollan and Frost (2001) have shown that response times to gender regular/transparent nouns were faster than the response times to gender opaque/irregulars forms. This effect was also observed in error rates. The authors observed that, in Hebrew, the phonological form (the ending of the nouns) of either regular or irregular nouns also influences the processing of grammatical gender decisions. Based on this finding, they argued that there are two routes for the processing and representation of grammatical gender in Hebrew: a lexically based route and a computational route. The lexically based route would be responsible for nouns whose grammatical gender is stored as information of the lemma of each noun. The computational route, on the other hand, would be responsible for the processing and representation of the nouns whose grammatical gender information can be recovered from their surface (phonological/orthographic) form.

Another very recent source of evidence in favor of a dual-mechanism for the processing of grammatical gender comes from Spanish. Afonso et al. (2013) employed a gender decision task and a masked priming task in order to assess two factors: first, whether words ending *-o*, typical of masculine nouns, and words ending in *-a*, typical of feminine nouns, are more relevant to gender assignment than other gender correlated endings such as *-ad* and *-ón*. For example, in Spanish, forms ending in *-ad* are associated with the feminine gender category (*ciudad* – “city”) and forms ending in *-ón* are associated with the masculine gender category (*embrión* – “embryo”, *bastión*– “stick”). Second, they investigated whether determiners are retrieved from the mental lexicon in conjunction with lexical items. In the gender assignment task, results clearly show that gender is easily assigned to masculine nouns ending in *-o*, and feminine nouns ending in *-a*, but participants were more prone to errors when the nouns were irregular, i.e., feminine nouns ending in *-o*, and masculine nouns ending in *-a*, and also when the words were opaque, that is, when the endings of the nouns were different from *-a* or *-o*. In addition, participants were faster to respond to regular forms than irregular or opaque forms. As regards the question of whether determiners facilitate gender assignment and whether they are retrieved together with the noun, the results of the masked priming task showed that participants were, again, faster in gender decisions for nouns ending in *-o* or *-a* than for nouns with other endings. Researchers observed that the definite article was relevant only when the nouns were opaque forms, suggesting

that the article is retrieved with the noun just for opaque forms that do not end in *-o* or *-a*. The authors interpreted these results as evidence for a dual-mechanism operating in gender assignment processing as well as in gender agreement processing.

Some recent neuroimaging studies have contributed to the dual route hypothesis as well. Hernandez *et al.* (2004, 2007) and Padovani *et al.* (2005) have found in a gender decision task that gender opaque forms show greater activation in Brodmann's area BA45 than gender transparent forms. Caffarra *et al.* (2015) have found different ERP components for gender transparent and gender opaque forms during sentence agreement processing.

The consistency of the results for different languages reveals that the processing of regular (transparent) forms seems to be quite different from the processing of opaque forms. These results seem to be analogous to results reported by studies on verbal morphology of English, which report differences between the processing of regular and irregular past tense formation. However, an issue that remains unresolved concerns the fact that while some studies report effects of frequency for gender transparent forms, other studies (TAFT; MEUNIER, 1998; DESROCHERS *et al.*, 1989) have shown that transparent/regular grammatical gender nouns can be stored.

Taft and Meunier (1998), for instance, indicated in a gender decision task that the phonological cues of French nouns play an important role in gender decisions. Participants were faster in regular nouns than in irregular ones. Nevertheless, effects of regularity and frequency also played a role in error rates and an interaction between regularity and frequency in reaction and error rates was not found. These results suggest that although phonological cues facilitate gender assignment, regular/transparent gender nouns can be stored in the lexical memory. These results, however, contradict those of a seminal study for French carried out by Tucker *et al.* (1968). French speakers were exposed to nonwords which had grammatical gender correlated endings. They found that the proportion of gender category (masculine or feminine) an item received was related to the proportion to which the ending was associated with the masculine or feminine gender category. Thus, their experiment showed that gender assignment to nonword was the product of phonological cues present in the end of the items and that gender assignment can be independent of lexical access.

In a gender assignment task, Bates *et al.* (1995) showed that participants were faster and more accurate in Italian nouns with highly predictable gender endings, thus providing evidence for effects of phonological cues as participants were faster in phonologically transparent nouns (masculine nouns ending in *-o*, and feminine nouns ending in *-a*) than in opaque nouns (other endings). Although the study controlled for frequency effects (high vs. low) and semantic features (abstractness vs. concreteness), results did not reveal any significance for these factors on the gender assignment task. Based on these findings, Bates *et al.* (1995) claim that there are two stages for gender assignment: First, the word is retrieved from the mental lexicon, and after that, a post-lexical check takes place in order to assign gender to the word. Faster reaction times to regular phonological cues are explained as convergence of information provided by the word ending and gender of the lexical item. Longer reaction times to irregular words are explained as a mismatch between the phonological cues and the grammatical gender.

In summary, the studies reviewed above show that phonological cues facilitate gender processing. Some researchers interpret this finding as the activation of a computational mechanism to process transparent grammatical gender lexical items. Nevertheless, other studies show that the frequency of items can influence gender processing, suggesting that memory plays a role independent of the form of the nouns. Thus, the debate remains open. Aiming to contribute to the discussion on the duality vs. singularity of grammatical gender processing, we carried out a behavioral study to investigate whether grammatical gender of transparent nouns and grammatical gender of opaque nouns in Brazilian Portuguese are processed by the same or by distinct cognitive mechanisms. In our study, we took into account the above mentioned frequency effects rationale, that is, if we find that frequency effects influence the processing of gender opaque nouns and do not influence the processing of transparent nouns, then our results suggest that regular/transparent gender nouns are computed while gender opaque forms are retrieved from memory. Based on the results encountered by Afonso *et al.* (2014) for Spanish, we hypothesize that the processing of grammatical gender in Portuguese is governed by a dual cognitive mechanism.

2 Grammatical gender in Portuguese

In Portuguese, every single noun has an intrinsic gender. Nouns in Portuguese can be either masculine or feminine. In most nouns in Portuguese, gender information can be retrieved from their phonological forms: nouns ending in *-a* are, generally, feminine, and nouns ending in *-o* are, generally, masculine. For instance, *menina* (“girl”) is a feminine noun and *menino* (“boy”) is a masculine noun. Similarly, the words *casa* (“house”) and *cama* (“bed”) are feminine nouns, whereas *carro* (“car”) and *corpo* (“body”) are masculine nouns, although they do not stand for an animate referent in the world. There are also other endings typical of feminine and masculine nouns. Examples of those endings are: nouns ending in *-agem* and *-ade* are, always, feminine nouns: *garagem* (“garage”), *imagem* (“image”), *idade* (“age”), *cidade* (“city”). Words ending in *-or* and *-ema* are, usually, masculine nouns: *amor* (“love”), *vapor* (“steam”), *problema* (“problem”), *cinema* (“cinema”), *sistema* (“system”).

It is important to emphasize that the general rule (feminine nouns ending in *-a*, and masculine nouns ending in *-o*) is the productive rule. Some studies for Brazilian Portuguese (e.g. NAME, 2002) have shown that children assign feminine gender to nonwords ending in *-a*, and masculine gender to nonwords ending in *-o*. This study also shows the same pattern for adults in gender assignment tasks. However, grammatical gender cannot always be predicted from a word’s endings. Actually, a huge amount of nouns are opaque as regards its grammatical gender. In other words, gender information cannot always be retrieved from a word’s phonological structure. In fact, in Portuguese there are nouns whose phonological form contradicts their gender, for instance, *dia* (“day”), *problema* (“problem”), *sofá* (“sofa”), *razão* (“sense”), *mão* (“hand”). As we can notice, although words such as *dia*, *problema*, and *sofá* end in *-a*, they are masculine, whereas words such as *mão* and *razão*, although ending in *-o* are feminine nouns. There are also other word endings which do not provide any cue for gender assignment such as *leite*_{masc}, *abacaxi*_{masc}, *noite*_{fem}, among many others.

Given this puzzling scenario, the question that arises is: How do native speakers of Portuguese know the gender of opaque forms (such as words as *leite*, *abacaxi*, *noite*) if gender assignment cannot be based on phonological cues? There are two possible answers to this question: The first answer could be that there are two different strategies

for gender agreement or gender assignment. This hypothesis relies on the fact that word endings function as a cue to gender agreement so that native speakers apply a very straightforward rule: if the word ends in *-a*, then it is feminine; if the word ends in *-o*, then it is a masculine noun. When such rule is not applicable, speakers rely on memory of the word's grammatical gender. This first answer implies that there is a dual route for gender assignment and gender agreement. The second answer is that grammatical gender assignment and agreement are based on the memory of the gender of each noun. Thus, native speakers of Portuguese simply know the gender of each noun. Therefore, gender assignment and agreement depend on memory. This answer implies that there is only one strategy for gender assignment and gender agreement and that there is no rule-governed behavior, which is in line with the single-mechanism account of language processing.

In the next section, we present the method employed in the present study to examine whether grammatical gender in Brazilian Portuguese is governed by a dual or by a single cognitive mechanism.

3 Method

In order to investigate whether a single or a dual route governs the processing of grammatical gender in Brazilian Portuguese, we employed a grammatical gender agreement task. Although studies on grammatical gender have traditionally used the gender assignment paradigm (e.g. TAFT; MEUNIER, 1998; AFONSO *et al.*, 2014; BATES *et al.*, 1995), we adopted the grammatical gender agreement task for two main reasons: first, because instead of gender assignment, the grammatical gender phenomenon is manifested in real language by means of gender agreement, i.e., in natural language, speakers make use of grammatical gender information for agreement purposes, not for gender assignment purposes. Second, since we intended to analyze the differences of grammatical gender processing between two conditions – article + noun (Condition 1) and noun + adjective (Condition 2) – it would not be possible to attribute gender to two words of different types such as noun and adjective at the same time since adjectives in Portuguese are either neutral (e.g. *feliz* – “happy”) or inflected by gender (e.g. *bonit-o*_{masc} or *bonit-a*_{fem} – “beautiful”).

The research question we address by conducting a grammatical gender agreement task is: Is the processing of grammatical gender governed by a single or a dual mechanism in i) the article + noun and ii) in the noun + adjective condition? We hypothesize that grammatical gender agreement processing is governed by a dual route and that we will find similar patterns in the article + noun agreement and in the noun + adjective agreement conditions. In both conditions, we do not expect to find frequency effects for the agreement of transparent forms but only for the opaque forms. A lack of frequency effects in transparent forms processing suggests that they are decomposed in their form in order to have gender agreement assignment. Instead of recurring to memory of the item's gender, speakers rely on the phonological cues, that is, the word's endings, to perform gender agreement both with the determiners and with the adjectives. On the other hand, since no phonological cues are provided by the opaque forms and the phonological cues of the opaque forms contradict their gender, speakers rely on memory of the item's gender to perform gender agreement. In what follows, we describe our task design.

3.1 Task Design

Materials: The entire set of stimuli comprised 112 items in total. We used the *Corpus Brasileiro* available online¹ to select 32 nouns (16 transparent and 16 opaque nouns) for condition 1: agreement between the definite article and the noun. We also selected 32 nouns (16 transparent and 16 opaque) to construct the stimuli for condition 2: agreement between the noun and the adjective. In this paper, we will refer to the items selected for condition 2 as *bigrams*. We considered transparent forms feminine words ending in *-a*, and masculine words ending in *-o*; we considered opaque forms all other word endings, including feminine words ending in *-o* and masculine words ending in *-a*. Half of the transparent nouns (8 items) and half of the opaque nouns (8 items) selected for condition 1 and condition 2 were high-frequent nouns (310-4608 per million) and the other half were low-frequent nouns (1-17 per million). We also tried to approximate the number of feminine and masculine nouns selected as much as possible. To compose condition 1, we selected seven masculine and nine feminine opaque nouns as well as nine masculine and seven

¹ www.linguateca.pt

feminine transparent nouns. To compose condition 2, we selected eight masculine and eight feminine opaque nouns as well as eight masculine and eight feminine transparent nouns.

The experiment also included 48 fillers (nonwords). The nonwords were phonologically legal letter strings. They were constructed by exchanging a vowel phoneme or a consonant phoneme of existing gender transparent and gender opaque nouns, e.g. **plefe* (correct form = *plebe*) or **visma* (correct form = *cisma*).

To construct the bigrams of condition 2, we looked up in the corpus the adjective most likely to co-occur with the noun selected. Besides manipulation of the frequency of the noun of a bigram, we also manipulated the frequency of occurrence in the corpus of the adjectives following the nouns in a bigram. The adjectives of the bigrams selected were either high-frequent or low-frequent adjectives so that we obtained frequent and infrequent nouns co-occurring with frequent and infrequent adjectives. As it was not always possible to find adjectives co-occurring with infrequent nouns in the *Corpus Brasileiro*, we used Google search to find adjectives likely to co-occur with the infrequent nouns selected. The adjectives of the bigrams selected from Google were also controlled by frequency in our analysis. The control of the frequency of the two elements of a bigram was adopted to investigate which one matters for the gender agreement: the frequency of the head of the noun phrase or of the adjective itself.

3.2 Apparatus

The experiment was programmed on *E-prime* software for psychological experiments. Items were presented in the center of the screen in a white lowercase size 18 against a black background of a LG Computer Monitor (Flatron L1755S).

3.3 Participants

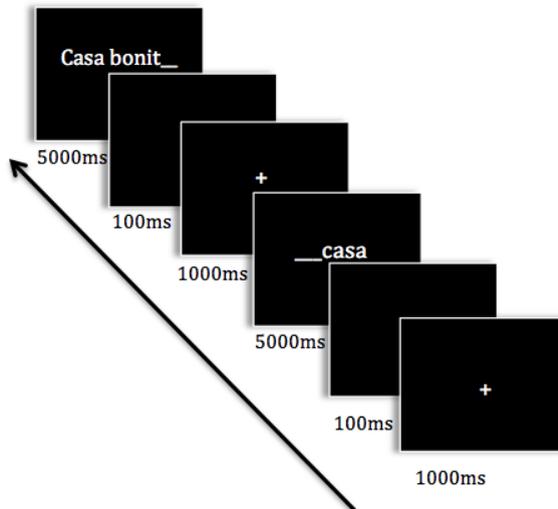
Participants were 19 right-handed native speakers of Brazilian Portuguese (9 female, $M_{Age} = 31$, range_{Age} = 21-44), recruited among graduate and undergraduate students at the Federal University of Santa Catarina (UFSC), Brazil. Subjects did not report previous register of neurological disorders. All participants reported having normal or corrected-to-normal vision. Subjects provided informed consent to

participate as volunteers in the present study. The study was approved by the research ethics committee of UFSC.

3.4 Procedure

The set of stimuli was presented visually and each subject performed the grammatical gender agreement task individually. The experimenter gave the task instructions to each participant while they appeared on the screen. The task consisted of a visual grammatical gender agreement in two conditions. Subjects were asked to push button 1 for letter *A* or button 5 for letter *O* which corresponded to the articles “A” and “O” preceding each noun or the ending of the adjectives to be inflected. Each trial was preceded by a fixation cross ‘+’ of 1000ms and the items remained on the screen up to 5000ms or until any button was pushed. We introduced 10 practice items before the beginning of each experimental session to make sure all participants were able to perform the main task properly. Figure 1 illustrates the presentation of the stimuli.

FIGURE 1 – Design of the experiment



NOTE: After the fixation cross, participants saw a blank screen followed by stimuli of condition 1 and condition 2. They then performed a gender agreement task. Stimuli were presented in different random order to each participant.

4 Results

4.1 Description of the data

Although we have coded participant's response for accuracy and reaction times, in this paper our analysis will be based on reaction times only. Inaccurate responses were removed from the dataset. We obtained 608 data points per condition, a total of 1216 data points. Table 1 displays description of reaction times collected for each category of condition 1 and condition 2.

TABLE 1 – Descriptive data per category

Conditions	Category	Mean	SD
Condition 1	Transparent frequent nouns	890ms	251.4
	Transparent infrequent nouns	1647ms	688.8
	Opaque frequent nouns	1033ms	304.8
	Opaque infrequent nouns	1184ms	462.4
Condition 2	Transparent frequent bigrams	1563ms	557.5
	Transparent infrequent bigrams	1549ms	582.8
	Opaque frequent bigrams	1413ms	611.7
	Opaque infrequent bigrams	1777ms	621.3

In condition 1, the results show that participants were faster to respond to transparent frequent nouns than to transparent infrequent nouns (890ms vs. 1647ms) $t = -14.243$, $df = 266.86$, $p\text{-value} < 2.2e-16$. Similarly, participants were faster to respond to opaque frequent nouns than to respond to opaque infrequent nouns (1033ms vs. 1184ms) $t = -2.9157$, $df = 283.25$, $p\text{-value} = 0.003$. These results show that frequency is a strong predictor for response speed regardless of the form of the nouns.

For condition 2, data description shows that participants were much faster to respond to opaque frequent bigrams than to respond to opaque infrequent bigrams (1413ms vs. 1777ms) $t = -5.8663$, $df = 295.75$, $p\text{-value} = 1.19$, suggesting that memory plays an important role in the processing of agreement between opaque forms and adjectives. On the other hand, reaction times for gender agreement between

frequent and infrequent transparent nouns and adjectives are very close (1549ms vs.1563ms) $t = 0.4102$, $df = 270.7$, $p\text{-value} = 0.68$, suggesting that a computational mechanism could be at work for agreement with transparent forms. Surprisingly, however, it is possible to notice that in condition 2 participants were slightly faster in agreement with transparent infrequent bigrams than in agreement with transparent frequent ones. Although the difference is statistically irrelevant ($p > 0.6$), these results might show that participants were faster in agreement with transparent infrequent bigrams because this category of items imposed more processing costs than agreement with transparent frequent bigrams, since nouns were first accessed in memory and, after that, were decomposed to perform agreement, while infrequent bigrams were only decomposed to perform agreement.

Since our task design contains two different conditions of gender agreement and each condition presents different types of factors that might affect reaction times, we adopted the *Linear Mixed Effects Models* approach with crossed effects of subjects and items for the statistical analysis to obtain a better picture of the data. The next section describes how the data analysis on reaction times was carried out.

4.1.1 Data analysis

The *Mixed Effects* models as well as data pre-processing were run in *Language R*. For the *Mixed Effects* models, we used the *lme4 package* (BATES, 2005; R Development Core Team, 2008). Prior to the data analysis, we screened the reaction times for extreme outliers ($RT < 400ms > 3000ms$). The outliers were removed from the data set of each category from condition 1 and condition 2 separately using *Grubbs test* for one or two outliers (GRUBBS, 1950), package *Outliers*. We removed 12 data points from condition 1 (1,97%) and 15 data points from condition 2 (2,4%). After cleaning the dataset, histograms showed that the distribution of reaction times of condition 1 and condition 2 were left-skewed. In order to minimize the skewness of the distribution, we log-transformed reaction times.

As will be shown in the next section, we ran separate models for condition 1 and condition 2. This procedure was adopted in order to investigate more closely the effects of predictors for each condition. The fixed effects we included in the models of condition 1 were: lemma

frequency of nouns (continuous), number of graphemes of nouns (continuous), sex/gender of participants (categorical), grammatical gender of the nouns, that is, whether the noun was feminine or masculine (categorical), number of graphemes of nouns (continuous) and number of phonemes of nouns (continuous). For condition 2, we included the following fixed effects: lemma frequency of the nouns (continuous), lemma frequency of adjectives (continuous), sex/gender of participants (categorical), grammatical gender of the nouns, that is, whether the noun of the bigram was feminine or masculine (categorical), and number of graphemes of bigrams (continuous). The number of graphemes of nouns and bigrams were included in the models as control factors in order to maximize statistical power by reducing residuals.

In each model, we included subjects and items (nouns for condition 1 and bigrams for condition 2) as random effects. In order to determine which random slope should be included in the models, we adopted a forward procedure, starting with a minimum model. Then, we included all possible random slopes for subject and items (words and bigrams) one by one. Models converged easily after scaling the continuous predictors. Following Baayen et al. (2008), model comparisons were made using likelihood ratio test to verify if the random slopes and interactions added to the minimum model were justified. As a result, the random slopes included were removed from the models of condition 1 and condition 2. P-values reported here were obtained by means of ANOVA type 3 which tests for the presence of significant interactions.

Table 2 displays the results of the model that best explains reaction times of gender agreement between article and nouns (condition 1). Table 3 exhibits results of the model that best explains reaction times of gender agreement between nouns and adjectives (condition 2). Table 1 and Table 2 show only factors that appear to have justifiable inclusion² in the final model according to results obtained after models comparison by means of likelihood ratio test.

² Following a standard procedure (see BARR *et al.*, 2013), we constructed various models for comparison purposes, using the likelihood ratio test. This test indicates the factors and interactions that are more likely to influence the dependent variable. Thus, the models presented in this paper do not always include all factors and interactions tested since they are the models that best fit our dataset.

TABLE 2 – Model that best explains reaction times for condition 1

Fixed effects	Estimate	Std.Error	t	Pr(>F)
Intercept	2.94986	0.06337	46.55	2.2e-16 ***
Lemma Frequency	0.00433	0.01456	0.30	0.7685
Category	0.02056	0.02028	1.01	0.3198
Number of graphemes	0.01499	0.01247	1.20	0.2402
Participants' Gender (Male)	0.02773	0.03472	0.80	0.4355
Nouns' Grammatical Gender	-0.02450	0.02111	-1.16	0.2563
Lemma Frequency:category	-0.14243	0.02338	-6.09	1.898e-06 ***
Random Effects				
Groups	name	Variance	Std. Deviation	Corr
subjects	intercept	0.005334	0.07303	
Word	Intercept	0.002329	0.04826	
Residual		0.011867	0.10893	

TABLE 3 – Model that best explains reaction times for condition 2

Fixed effects	Estimate	Std.Error	t	Pr(>F)
Intercept	3.253e+00	1.060e-01	30.695	0.002588 **
Lemma Frequency (nouns)	-6.614e-06	1.985e-06	-3.332	0.002588 **
Category	-4.639e-02	2.896e-02	0.61	0.121367
Number of graphemes	-5.688e-03	9.800e-03	-0.58	0.568410
Participants' Gender	2.077e-02	4.111e-02	0.50	0.619894
Lemma Frequency (nouns): Category	5.949e-06	2.303e-06	2.583	0.015618 *
Grammatical gender	2.338e-02	2.187e-02	1.069	0.295044
Random Effects				
Groups		Variance	Std. Deviation	corr
subjects	Intercept	0.007654	0.08749	
Bigram	Intercept	0.003083	0.05553	
Residual		0.010900	0.10440	

The model of condition 1 shows a significant interaction between category (transparent nouns vs. opaque nouns) with frequency (high vs. low). This interaction means that frequency affects categories of nouns differently. The interaction plot of Figure 2 illustrates the effect of frequency on reaction times of gender agreement between articles and nouns. Importantly, frequency of lemma itself and category itself did not reach significance. Factors such as number of graphemes and grammatical gender of the nouns (masculine vs. feminine) did not influence reaction times. Participants' gender (male vs. female) did not influence reaction times. We did not find a three-way interaction of gender, frequency, and category ($\beta=0.010096$, $t=0.53$), which suggests that reaction times of both men and women were equally influenced by interaction between frequency and category of nouns. We did not find a three-way interaction of grammatical gender, frequency, and category ($\beta=4.528e-02$, $t=0.55$) either, suggesting that reaction times for both masculine and feminine nouns were equally influenced by interaction between frequency and category of nouns.

As regards condition 2, results show a main effect of lemma frequency of nouns and a significant interaction between lemma frequency of nouns and category. As frequency of adjectives did not reach significance neither did a three-way interaction of lemma frequency of nouns, lemma frequency of adjectives, and category, our final model includes only the interaction between lemma frequency of nouns and category. As shown in Table 3, similar to condition 1, there is an interaction between lemma frequency of nouns and category, indicating that lemma frequency of nouns influenced reactions times of transparent and opaque bigrams differently. However, the effect of lemma frequency in condition 2 is more prominent than in condition 1, especially on opaque forms. The interaction plot of Figure 3 illustrates this effect. Concerning the other predictors included in our final model of condition 2 (participants' gender, grammatical gender of nouns, and number of graphemes of bigrams), as in condition 1, those predictors failed to reach significance as revealed by p-values presented in Table 3.

FIGURE 2 – Interaction Plot Condition 1 – Transparent forms x Opaque forms

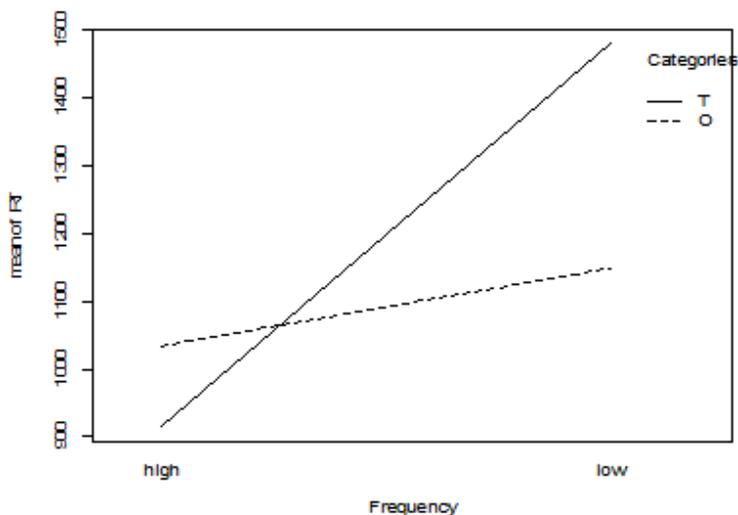
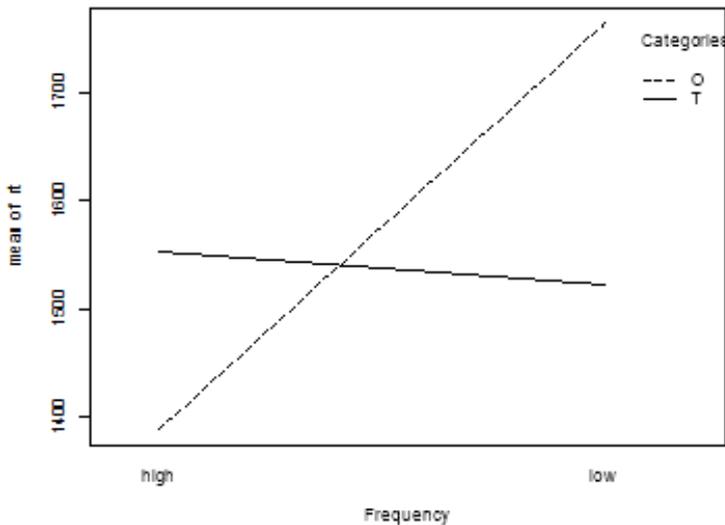


FIGURE 3 – Interaction Plot Condition 2 – Transparent forms x Opaque forms



5 Discussion

In the present study, we investigated the processing routes of grammatical gender agreement in Brazilian Portuguese in two conditions: between an article and a noun (condition 1) and between a noun and an adjective (condition 2). Results suggest that grammatical gender agreement processing between article and noun differs from that between noun and adjective. We can notice, by observing reaction times (see Table 1), that agreement between noun and adjectives elicits more processing costs, as reflected by longer reaction times, than agreement between article and nouns. It is our contention that, as subjects had more information to process in condition 2 than in condition 1 (for instance, a higher number of phonemes and graphemes, two items for retrieval, items decomposition, and concatenation), reaction times were longer as a result of the fact that more information demands more time to be processed.

In condition 1, while an interaction between frequency and form category plays a role in reaction times, frequency and form category alone fail to reach significance. This result suggests that form category itself is not a predictor of reaction times and neither is frequency itself. In fact,

there is a combination of both factors, but with differences concerning the influence of frequency on reaction times between form categories. This is very clear when we compare the reaction times of frequent transparent forms and frequent opaque forms. Participants were faster in gender agreement for frequent transparent forms compared to frequent opaque forms, infrequent transparent forms and infrequent opaque forms, replicating previous findings for other languages such as Italian (BATES et al., 1995), French (TAFT; MEUNIER, 1998; DESROCHERS et al., 1989; TUCKER et al., 1968), German (HOHFELD, 2006), Hebrew (GOLLAN; FROST, 2001) and Spanish (AFONSO et al., 2013). The faster reaction times to respond to frequent transparent forms than to respond to frequent opaque forms suggest that the phonological cues of the transparent forms yield a processing advantage as compared to a processing disadvantage for gender agreement with opaque forms, which, in turn, suggests that phonological cues facilitate gender processing.

Although in condition 1 our results indicate that high-frequency transparent forms can be stored, replicating results of Taft and Meunier (1998) and Desrochers et al. (1989), our findings suggest that low-frequency items might be decomposed. One of the clues to support this assumption is that, in condition 1, it is possible to observe that participants were slower to process infrequent transparent forms than infrequent opaque forms. A possible explanation for this phenomenon would be an increase in the processing costs while decomposing transparent infrequent nouns to retrieve gender information. This finding suggests that, on the one hand, high-frequency transparent forms are fully accessed in memory, but infrequent items might be decomposed. Thus, results of condition 1 suggest that computation would be applicable only when access to items in memory is difficult. Our assumption is based on the fact that other studies (e.g. Laine, Vainio and Hyönä, 1999) have shown that decomposition and concatenation processes are more demanding than lexical access, resulting in longer reaction times.

In condition 2, results also show that frequency of nouns plays a role in reaction times. As in condition 1, our final model shows a significant interaction between lemma frequency of the nouns and category, suggesting that frequency influences word categories differently. However, in condition 2, as illustrated by the interaction plot of Figure 3, it is possible to observe that the influence of lemma frequency of the nouns is stronger on opaque forms than on transparent forms, indicating

that opaque forms are stored and transparent forms might be computed. As previously mentioned, surprisingly, subjects were slightly slower to perform gender agreement with high-frequent transparent forms than with low-frequent transparent forms. Following Bates et al. (1995) rationale, we maintain that this phenomenon might be related to a three-stage processing strategy for high-frequent transparent forms – first lexical access, second item decomposition, third affix concatenation – as opposed to a two-stage processing for the low-frequency items (decomposition followed by affix concatenation). How to interpret these results in the light of the dual mechanism model and single-mechanisms model?

One may argue that since we found different patterns for both types of phonological forms, a dual mechanism would be at work in the processing of grammatical gender between an article and a noun and a noun and an adjective in Portuguese. However, we claim, on the basis of the frequency effects that were observed influencing the processing of both transparent and opaque forms, that there is a single mechanism governing grammatical gender processing in condition 1. As regards condition 2, in our view, there are two mechanisms governing gender agreement processing between nouns and adjectives. However, our results also indicate that, even if a computational processing strategy is operating, items are still being lexically accessed. This interpretation is based on the slightly longer reaction times in high-frequent transparent nouns in condition 2, suggesting an *a priori* lexical access and further decomposition and concatenation, leading to longer reaction times.

Our experiment shows that both transparent and opaque gender forms can be stored, but computational processing can be activated under some circumstances: when subjects do not readily extract the items from the mental lexicon and, as in condition 2, when there is more information to be processed. It seems that when the strings are longer as they are in condition 2, subjects prefer to decompose them to perform agreement. They also prefer to decompose items that are not easily accessed in memory as occurred with infrequent gender transparent nouns in condition 1. As observed by Rastle and Davis (2008), although more time-consuming, decomposition is an easier strategy when items present clues on their surface form.

Our results indicate that memory is involved regardless of the form of the noun when it is a frequent noun, that is, subjects recognize

the item and look up for it in the mental lexicon. Thus, we suggest that it is not the case that a computational route is activated every time subjects face a transparent form. The first strategy is to search for the item in the mental lexicon. If the item cannot be found, a second strategy takes place: decomposition. Based on these findings, we argue that our results do not support the dual mechanism view of grammatical gender processing and, therefore, do not support our initial hypothesis that the processing of grammatical gender in Portuguese is governed by a dual cognitive mechanism. On the contrary, our experiments show that regular/transparent forms can be stored as whole words, as predicted by single mechanism models.

It is noteworthy that the number of participants selected and number of items selected limited the experiment presented here. An experiment containing a larger sample of participants and items could provide different results. Future research with more subjects and more levels of items frequency is necessary to verify whether very infrequent transparent nouns are in fact subject to decomposition. In order to disentangle this issue, it would be of great aid to test if medium-frequency gender transparent nouns are also subject to decomposition or if they can also be accessed in memory just as high-frequency gender transparent nouns.

Acknowledgement

This research was supported by CAPES/NUFFIC Grant 051/13 – 23038.007129/2013-40.

References

- AFONSO, O.; DOMÍNGUEZ, A.; ÁLVAREZ, C. J.; MORALES, D. Sublexical and lexico-syntactic factors in gender access in Spanish. *Journal of Psycholinguistic Research*, Springer Link, v. 43, Issue 1, p. 13-25, Feb. 2014.
- ALEGRE, M.; GORDON, P. Frequency Effects and the representational status of regular inflections. *Journal of Memory and Language*, Elsevier, v. 40, p. 41-61, 1999. <https://doi.org/10.1006/jmla.1998.2607>.

BAAYENS, H.; DIJKSTRA, T.; SCHREUDER, R. Singulars and Plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, Elsevier, v. 37, p. 94-117, 1997. <https://doi.org/10.1006/jmla.1997.2509>.

BAAYEN, R. H.; DAVIDSON, D. J.; BATES, D. M. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, Elsevier, v. 59, p. 390-412, 2008. <https://doi.org/10.1016/j.jml.2007.12.005>.

BAAYEN, H. R., TWEEDIE, F. J., SCHREUDER, R. The subjects as a simple random effect fallacy: Subject variability and morphological family effects in the mental lexicon, *Brain and Language*, v. 81, p. 55-65, 2002. <https://doi.org/10.1006/brln.2001.2506>.

BARR, D.J.; LEVY, R.; SCHEEPERS, C.; TILY, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, Elsevier, v. 68, p. 255-278, 2013. <https://doi.org/10.1016/j.jml.2012.11.001>.

BATES, E.; DEVESCOVI, A.; PIZZAMIGLIO, L.; D'AMICO, S.; HERNADEZ, A. Gender and lexical access in Italian. *Perception & Psychophysics*, ResearchGate, v. 57, p. 847-862, 1995.

BOWDEN, H. W. *Proficiency and second-language neurocognition: a study of Spanish as a first and second language*. 2007. 241f. Tese (Doutorado) – Georgetown University, Washington D.C., 2007.

BYBEE, J. L. *Morphology: a study of the relation between meaning and form*. Philadelphia: John Benjamins, 1985. <https://doi.org/10.1075/tsl.9>.

BYBEE, J. L. Regular Morphology and the Lexicon. *Language and Cognitive Processes*, Taylor & Francis Online, v. 10, p. 425-455, 1995.

CAFFARRA, S., SYIANOVA-CHANTURIA, A., PESCIARELLI, F., VESPIGNANI, F., CACCIARI, C. Is the noun ending a cue to grammatical gender processing? An ERP study on sentences in Italian. *Psychophysiology*, Wiley Online Library, v. 52, n. 8, p. 1019-1030, 2015.

CHOMSKY, Noam. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CLAHSEN, H.; ROTHWEILER, M.; WOEST, A. Regular and irregular in the acquisition of German noun plurals. *Cognition*, Elsevier, v. 45, p. 225-255, 1992. [https://doi.org/10.1016/0010-0277\(92\)90018-D](https://doi.org/10.1016/0010-0277(92)90018-D).

CLAHSEN, H.; EISENBEISS, S.; SONNENSTUHL, I. Morphological structure and the processing of inflected words. *Theoretical Linguistics*, De Gruyter, v. 23, p. 201-249, 1997.

CORBETT, G. *Gender*. Nova York: Cambridge University Press, 1991.

DESROCHERS, A.; PAIVIO, A.; DESROCHERS, S. L'effet de la fréquence d'usage des noms inanimés et de la valeur prédictive de leur terminaison sur l'identification du genre grammatical. *Revue Canadienne de Psychologie*, Société Canadienne de Psychologie, v. 43, p. 62-73, 1989. <https://doi.org/10.1037/h0084253>.

FLEISCHHAUER, E.; CLAHSEN, H. Generating Inflected Word Forms in Real Time: Evaluating the Role of Age, Frequency, and Working Memory. In: ANNUAL BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT, 36th., Somerville, 2012. *Proceedings...* Somerville: Cascadilla Press, 2012.

GOLLAN, T.; FROST, R. Two routes to grammatical gender: Evidence from Hebrew. *Journal of Psycholinguistic Research*, Springer Lind, v. 30, n. 6, p. 627-651, 2001.

GRUBBS, F. Procedures for Detecting Outlying Observations in Samples. *Technometrics*, Taylor & Francis Online, v. 11, n. 1, p. 1-21, 1969.

HERNANDEZ, A. E.; HOFMANN, J.; KOTZ, S. A. Age of acquisition modulates neural activity for both regular and irregular syntactic functions. *NeuroImage*, Elsevier, v. 36, p. 912-923, 2007. <https://doi.org/10.1016/j.neuroimage.2007.02.055>.

HERNANDEZ A. E.; KORTZ A. S.; HOFMANN J.; VALENTIN V. V., DAPRETTO M.; BOOKHEIMER S. The neural correlates of grammatical gender decisions in Spanish. *Neuroreport*, LWW Journals, v. 15, n. 5, p. 863-866, 2004.

HOHFELD, A. Accessing grammatical gender in German: the impact of gender-marking regularities. *Applied Psycholinguistics*, Cambridge University Press, v. 27, p. 127-142, 2006.

JACKENDOFF, R. What's in the lexicon?. In: NOOTEBOOM, S.; WEERMAN, F.; WIJNEN, F. (Org.). *Storage and Computation in the Language Faculty*. Dordrecht: Kluwer Academic Publishers, 2002. p. 61-92. https://doi.org/10.1007/978-94-010-0355-1_2.

LAINE, M.; VIANIO, S.; HYÖDNÄ, J. Lexical access routes to nouns in a morphologically rich language. *Journal of Memory & Language*, Elsevier, v.40, p.109-135, 1999. <https://doi.org/10.1006/jmla.1998.2615>.

MACWHINNEY, B. The competition model: the input, the context, and the brain. In: ROBINSON, P. (Org.). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, 2001. p. 69-90. <https://doi.org/10.1017/CBO9781139524780.005>.

MARCUS, G. F.; BRINKMANN, U.; CLAHSSEN, H.; WIESE, R.; PINKER, S. German inflection: The exception that proves the rule. *Cognitive Psychology*, Elsevier, n. 29, p. 189-256, 1995.

MORGAN-SHORT, K.; SANZ, C.; STEINHAEUER, K.; ULLMAN, M.T; second language acquisition of gender agreement in explicit and implicit training conditions: an event-related potential study. *Language Learning*, Wiley Online Library, v. 60, 1, 154-193, 2010. <https://doi.org/10.1111/j.1467-9922.2009.00554.x>.

NAME, M. C. L. *Habilidades perceptuais e linguísticas no processo de aquisição do sistema de gênero no português*. Ph.D. Dissertation. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro, 2002.

NEWMAN, A. J.; ULLMAN, M. T.; PANCHEVA, R.; WALIGURA, D. L.; NEVILLE, H. J. An ERP study of regular and irregular English past tense inflection. *NeuroImage*, Elsevier, v. 34, p. 435-445, 2007. <https://doi.org/10.1016/j.neuroimage.2006.09.007>.

PADOVANNI, R.; CALANDRA-BUONAURO, G.; CACCIARI, C.; BENUZZI, F.; NICHELLI, P. Gramatical gender in the brain: Evidence from an fMRI study on Italian. *Brain Research Bulletin*, Elsevier, v. 65, p. 301-308, 2005. <https://doi.org/10.1016/j.brainresbull.2004.11.025>.

PINKER, S. Rules of Language. *Science*, American Association for the Advancement of Science, v. 253, p. 530-535, 1991. <https://doi.org/10.1126/science.1857983>.

PINKER, S. Words and Rules. *Lingua*, Elsevier, v. 106, p.219-242, 1998. [https://doi.org/10.1016/S0024-3841\(98\)00035-7](https://doi.org/10.1016/S0024-3841(98)00035-7).

PINKER, Steven. *Words and rules: the ingredients of Language*. Londres: Weidenfeld and Nicolson, 1999.

PRADO, E.; ULLMAN, M. T. Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology: Language, Memory and Cognition*, v. 35, n.4, p.849-866, 2009. <https://doi.org/10.1037/a0015286>.

PRASADA, S.; PINKER, S. Generalizations of regular and irregular morphology. *Language and Cognitive Processes*, Taylor & Francis Online, v. 8, n. 1, p. 1-56, 1993.

RASTLE, K.; DAVIS, M. Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, Taylor & Francis Online, v. 23, p. 942-971, 2008.

R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-9000051-07-0, URL: <http://www-R-project.org>, 2008.

TAFT, M.; MEUNIER, F. Lexical representation of gender: A quasiregular domain. *Journal of Psycholinguistic Research*, Springer Link, v. 27, n. 1, p. 23-45, 1998.

TUCKER, G. R.; LAMBERT, W. E.; RIGAUULT, A. A.; SEGALOWITZ, N. A psychological investigation of French speaker's skill with grammatical gender. *Journal of Verbal Learning and Verbal Behavior*, ResearchGate, v. 7, p. 312-316, 1968.

ULLMAN, M.T. Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighbourhood effects. *Language and Cognitive Processes*, Taylor & Francis Online, v. 1, n. 14, p. 47-67, 1999. <https://doi.org/10.1080/016909699386374>.

ULLMAN, M.T. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, Springer Link, v. 30, n. 1, p. 37-69, 2001.

ULLMAN, M.T. Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *Cognition*, v. 92, n. 1-2, p. 231-270, 2004. <https://doi.org/10.1016/j.cognition.2003.10.008>.

ULLMAN, M.T. A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model. In: SANZ, C. (Org). *Adult second language acquisition*. Washington, DC: Georgetown University Press, 2005. p. 141-178.

VAN DER LELY, H. K. J.; ULLMAN, M. T. Past tense morphology in specifically language impaired and normally developing children. *Language and Cognitive Processes*, Taylor & Francis Online, v. 16, n. 2/3, p. 177-217, 2001.

Appendix

TABLE 1 – Stimuli list of condition 1

Item	Grammatical gender	Lemma Frequency (noun)	Form Category
__ dia	masc	2119	Opaque
__ saúde	fem	320	Opaque
__ clave	fem	1	Opaque
__ coldre	masc	1	Opaque
__ covil	masc	1	Opaque
__ dogma	masc	3	Opaque
__ filme	masc	597	Opaque
__ gente	fem	822	Opaque
__ lei	fem	572	Opaque
__ limiar	masc	6	Opaque
__ mão	fem	905	Opaque
__ noite	fem	616	Opaque
__ ode	fem	1	Opaque
__ país	masc	1461	Opaque
__ plebe	fem	1	Opaque
__ prole	fem	3	Opaque
__ prelo	masc	2	Transparent
__ ata	fem	10	Transparent
__ átrio	masc	1	Transparent
__ bojo	masc	5	Transparent
__ casa	fem	1062	Transparent
__ casta	fem	7	Transparent
__ fleuma	fem	1	Transparent
__ hora	masc	480	Transparent
__ hulha	fem	1	Transparent
__ jogo	masc	689	Transparent
__ jugo	masc	2	Transparent
__ livro	masc	648	Transparent
__ pessoa	fem	385	Transparent
__ preço	masc	407	Transparent
__ tempo	masc	1515	Transparent
__ vida	fem	1059	Transparent

TABLE 2 – Stimuli list of condition 2

Items	Grammatical gender	Form Category	Lemma Frequency (noun)	Lemma Frequency (Adj)
feixe rápid__	Masc	Opaque	7	291
lider inat__	Masc	Opaque	440	4
base sólíd__	Fem	Opaque	417	46
time barat__	Masc	Opaque	652	103
rede comprid__	Fem	Opaque	411	28
prece calm__	Fem	Opaque	8	27
fronte baix__	Fem	Opaque	15	433
morte súbit__	Fem	Opaque	486	34
motim louc__	Masc	Opaque	7	68
série nov__	Fem	Opaque	310	2370
nuance nov__	Fem	Opaque	4	2370
haste comprid__	Fem	Opaque	5	28
âmbar clar__	Masc	Opaque	1	428
matiz escur__	Masc	Opaque	6	69
papel gross__	Masc	Opaque	468	52
total alt__	Masc	Opaque	410	1014
vulto pret__	Masc	Transparent	17	148
mundo inteir__	Masc	Transparent	1076	185
Água parad__	Fem	Transparent	421	43
rua escondid__	Fem	Transparent	612	6
gueto escur__	Masc	Transparent	10	69
fava branc__	Fem	Transparent	5	308
beca antig__	Fem	Transparent	1	318
banco privad__	Masc	Transparent	720	299
crina long__	Fem	Transparent	1	427
coisa cert__	Fem	Transparent	1222	700
obra acabad__	Fem	Transparent	722	5
ano letiv__	Masc	Transparent	4608	6
gnomo velh__	Masc	Transparent	1	414
prazo long__	Masc	Transparent	532	427
pompa suntuos__	Fem	Transparent	5	3
clero antig__	Masc	Transparent	4	318

A influência da referencialidade no processamento de orações relativas em português brasileiro

The influence of referentiality on relative clause processing in Brazilian Portuguese

Gitanna Brito Bezerra

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil
gitannabezerra@gmail.com

Márcio Martins Leitão

Universidade Federal da Paraíba, João Pessoa, Paraíba / Brasil
profleitao@gmail.com

Lorrane da Silva Neves Medeiros

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
lorranesnm@hotmail.com

Resumo: Neste artigo apresentamos dois experimentos que foram conduzidos para investigar, no português brasileiro, o papel da referencialidade no processamento de orações relativas associadas a Sintagmas Nominais complexos (SNs complexos) do tipo substância (como “a bolsa de (do) couro”). O experimento de leitura automonitorada revelou uma preferência imediata pela aposição não local, evidenciando uma forte influência da referencialidade do N1 no processamento da oração relativa. O experimento de rastreamento ocular, por sua vez, revelou um efeito imediato da referencialidade do N2: uma aposição local é favorecida quando o N2 é referencial e uma aposição não local é preferida quando o N2 não é referencial. Nossos dados vão na direção do que é proposto pela Hipótese de *Construal* (FRAZIER; CLIFTON, 1996).

Palavras-chave: processamento; orações relativas; referencialidade.

Abstract: In this paper we report two experiments that were conducted to investigate the influence of referentiality on the processing of relative clauses associated with complex noun phrases of substance reading (like “the bag of (the) leather”). The self-paced reading task showed an immediate preference for high-attachment, revealing a strong influence of N1 referentiality on relative clause processing. The eyetracking study, in turn, revealed an immediate effect of N2 referentiality: low attachment is favored when the N2 is referential and high attachment is preferred when the N2 is non-referential. Our data are in tune with the general proposal of the Construal Hypothesis (FRAZIER; CLIFTON, 1996).

Keywords: processing; relative clauses; referentiality.

Recebido em: 9 de dezembro de 2016.

Aprovado em: 31 de janeiro de 2017.

1 Introdução

Na área da Psicolinguística Experimental, diferentes modelos de processamento sentencial buscam explicitar a maneira pela qual os sujeitos compreendem e produzem linguagem. No que tange, especificamente, ao processo de compreensão linguística, tais modelos diferem-se, de forma geral, quanto ao caminho proposto e assumido para se alcançar a interpretação de uma sentença. Nesse âmbito, a Teoria do *Garden Path* (FRAZIER, 1979; FRAZIER; RAYNER, 1982) propõe que a interpretação de uma sentença pressupõe uma estrutura sintática: o significado de uma sentença vai sendo construído composicionalmente à medida que o *parser*, o processador sentencial, atribui um esqueleto sintático ao *input* linguístico. Essa teoria assume, mais especificamente, que o *parser*, na atribuição inicial de uma representação sintática para uma cadeia linguística, tem acesso restrito à informação sintática, não podendo, por exemplo, acessar de imediato informações discursivas para decidir se apõe um sintagma X a um local Y ou a um local Z. Decisões desse tipo, oriundas de ambiguidades estruturais, seriam, na verdade, automaticamente feitas seguindo um princípio geral: escolha a primeira análise disponível (FRAZIER, 1987). Uma aposição é imediatamente determinada, considerando, primeiramente, a análise que pedir o menor número de nós sintáticos (Aposição Mínima) e, no caso de duas análises

com mesma complexidade estrutural, a análise que implicar uma aposição ao sintagma ou oração correntemente sendo processado (Aposição Local).

Essa necessária determinação inicial de análise foi repensada quando a Teoria do *Garden Path* foi reformulada, dando origem à Hipótese de *Construal* (FRAZIER; CLIFTON, 1996). Uma das motivações para essa reformulação foi o surgimento de evidências experimentais divergentes quanto às predições do princípio da Aposição Local. Em uma sentença como “Alguém atirou no empregado da atriz que estava na varanda”, tal princípio prediz uma preferência sistemática pela aposição da oração relativa ao N2, “a atriz”, e não ao N1, “o empregado”. Porém, em um trabalho seminal, Cuetos e Mitchell (1988) reportaram evidências de que falantes de Espanhol preferiram, na verdade, a aposição ao N1, enquanto falantes de Inglês preferiram a aposição ao N2. A partir de então, este tipo de estrutura com oração relativa passou a ser foco de estudos experimentais em diversas línguas, originando-se diferentes hipóteses na tentativa de explicar a ausência de uma sistematicidade nas preferências de análise entre as línguas (CUETOS; MITCHELL, 1988; GIBSON; PEARLMUTTER; CANSECO-GONZALEZ; HICKOK, 1996; FODOR, 1998, 2002; HEMFORTH; KONIECZNY; SEELIG; WALTER, 2000; MAIA; FERNÁNDEZ; COSTA; LOURENÇO-GOMES, 2006; GRILLO; COSTA, 2014; HEMFORTH; FERNANDEZ; CLIFTON; FRAZIER; KONIECZNY; WALTER, 2015).

Nesse contexto, o *Construal* surgiu reconhecendo a pertinência do princípio da Aposição Local e enfatizando que, antes de dar ênfase às diferenças de aposição observadas entre as línguas e pôr em cheque a universalidade de tal princípio, seria preciso observar as diferentes preferências de análise existentes em uma mesma língua a depender do tipo de estrutura focalizado. Assim, o *Construal* formalizou uma distinção entre sintagmas primários e sintagmas secundários. O primeiro tipo diz respeito ao sujeito e ao predicado principal de uma oração finita, bem como aos seus complementos e constituintes obrigatórios. O segundo tipo diz respeito aos sintagmas que não podem, nem mesmo temporariamente, ser analisados como sintagmas primários. Os sintagmas secundários ditos “potencialmente primários” terão seu status “primário” acessado em primeiro lugar. Simplificando, o primeiro tipo de relação alude a sintagmas argumentos e o segundo alude a sintagmas adjuntos. A chave dessa distinção é a assunção de que o *parser* emprega mecanismos de análise distintos dependendo do tipo de estrutura em questão. Dessa

maneira, diante de uma ambiguidade envolvendo a análise de um sintagma primário, o *parser* está sujeito a uma pressão para determinar a descrição lexical de um dado núcleo lexical e procede a uma aposição imediata seguindo princípios estruturais – uma análise determinada, então, é automaticamente construída. No entanto, tal pressão não está presente em uma ambiguidade que envolve um adjunto, dado que esse elemento não está envolvido na especificação das propriedades de núcleos lexicais, dando oportunidade para uma indeterminação ou subespecificação inicial de análise – o *parser*, nesse caso, procede a uma associação.

Com essa distinção feita, Frazier e Clifton (1996) lançaram a hipótese de que o *parsing* de orações relativas do tipo mencionado anteriormente não exibiria uma sistemática preferência de análise entre as línguas pois ele estaria sujeito à influência não só de informações estruturais, mas também, e crucialmente, de informações não-estruturais. Investigando os fatores que poderiam influenciar o processamento das relativas, os autores propuseram, por exemplo, o Princípio de Referencialidade, segundo o qual orações relativas restritivas são preferencialmente apostas a núcleos referenciais. É este princípio que a presente pesquisa focaliza.

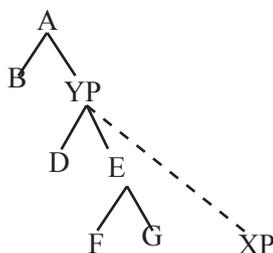
Nós reportaremos dois experimentos em português brasileiro (uma leitura automonitorada e um rastreamento ocular) que realizamos para investigar a influência desta informação não-estrutural de referencialidade. Considerando os resultados obtidos, discutiremos até que ponto os sujeitos deixam-se guiar por informações não-estruturais no processamento *on-line* e até que ponto eles são sensíveis e fiéis à informação de gênero como material desambiguador. Para embasarmos nossa discussão, organizamos o texto da seguinte forma: no tópico 2, abordaremos a proposta do *Construal*, apresentando a noção de subespecificação inicial de análise e o princípio da Referencialidade; no tópico 3, apresentaremos os dois experimentos realizados; e, no tópico 4, apresentamos uma discussão geral dos resultados obtidos.

2 *Construal* de orações relativas

A Hipótese do *Construal* (FRAZIER; CLIFTON, 1996, 1997) para o processamento de sintagmas secundários é a seguinte: associe um sintagma secundário ao domínio de processamento temático (a projeção

máxima estendida do último atribuidor temático) e o interprete usando informações estruturais e não estruturais. O processo de associação, especificamente, pressupõe a possibilidade de uma subespecificação limitada de análise, já brevemente mencionada, a qual alude ao fato de que, na construção de uma representação sintática, o *parser* pode deixar relações sintáticas de dominância e irmandade não completamente especificadas. Na figura a seguir, essa indeterminação inicial na análise é evidenciada pela linha tracejada que associa o sintagma secundário XP à estrutura através do YP, que, neste caso, seria a projeção máxima do último atribuidor temático:

FIGURA 1 – Associação de um sintagma secundário



Fonte: Frazier; Clifton (1997, p. 280).

Embora os autores defendam que o *parser* pode não determinar uma análise estrutural de imediato, a proposição de uma associação inicial a um domínio temático específico garante que o sintagma secundário receba uma estruturação mínima (não ficando, portanto, totalmente sem uma estruturação, “solto”, o que, se fosse o caso, afetaria a sua manutenção na memória de trabalho) até que, no processo posterior de interpretação, receba uma análise específica. Considerando a figura 1, no processo de interpretação, o acesso a informações sintáticas e não sintáticas permite que um local determinado de aposição seja selecionado para o sintagma secundário (XP): “*If XP associates to YP then XP must take YP, or some node dominated by YP, as its sister*” (FRAZIER; CLIFTON, 1997, p. 279). Entende-se, portanto, que, depois da subespecificação inicial, um compromisso sintático é eventualmente realizado, alcançando-se uma representação estrutural com as relações sintáticas devidamente determinadas.

No tocante, especificamente, aos processos de associação e de interpretação de uma oração relativa, uma informação não estrutural que, no escopo do *Construal*, é tratada como importante no processo de determinação de análise é a referencialidade do N1 e do N2. De maneira geral, o *status* referencial de um dado nome relaciona-se ao seu potencial ou não de referir a uma entidade discursiva presente em um modelo discursivo ou de introduzir uma entidade discursiva no modelo discursivo em construção para uma sentença. Com base nesta noção do que seria um nome referencial, Gilboy, Sopena, Clifton e Frazier (1995, p. 136) propõem o Princípio de Referencialidade:

the heads of some maximal projections are referential in the sense that they introduce discourse entities (e.g., participants in events described in the discourse) into a discourse model (at least temporarily), or correspond to already existing discourse entities. Restrictive modifiers (e.g., restrictive relative clauses) preferentially seek hosts which are referential in this sense.

Em específico, os autores assumem que um nome é referencial quando ele é introduzido por um determinante. Seguindo o princípio, se houver, no domínio temático corrente, um nome referencial e um nome não referencial, a relativa será aposta ao nome que é referencial, tendo em vista que ele, correspondendo a uma entidade discursiva, estará disponível para uma modificação restritiva, justificando, assim, a função da oração relativa. Veja-se que este princípio não teria uma predição específica para casos envolvendo sintagmas nominais como “a filha do coronel”, já que ambos os nomes se encontram no mesmo domínio temático e ambos são referenciais. O princípio, conforme os autores, poderia prever uma preferência específica para sintagmas nominais como “a mesa de madeira” e “a taça de vinho”, sendo o primeiro do tipo “substância” ou “material” e, o segundo, do tipo “quantidade” (ou ainda “conteúdo” e “classificação”). Nesses tipos de sintagmas, é mais natural que o N2 apareça sem um determinante, não sendo, portanto, referencial, sendo todo o SN complexo visto como uma única entidade discursiva. Por consequência, nestes casos, o N1, que é referencial, é visto como o local preferido de aposição da oração relativa.

Gilboy *et al.* (1995) investigaram, dentre outros fatores, esta influência do *status* referencial no processamento de orações relativas

através de dois questionários. Focalizando aqui o SN “substância”, foi encontrada, em questionários com falantes de inglês e falantes de espanhol uma preferência geral pela aposição da oração relativa ao N1 em sentenças como “*Yesterday they gave me the sweater of cotton that was illegally imported*” e (versão em espanhol: “*Ayer me regalaron el jersey de algodón que importaban de contrabando*”), as quais contêm apenas um nome referencial: o N1. Em um questionário posterior, agora só com falantes de inglês, os autores manipularam a referencialidade do N2, tornando-o também referencial: “*Yesterday they gave me the sweater of the cotton that was illegally imported*”. A predição era a de que a quantidade de respostas N2 aumentaria neste caso, tendo em vista que o N2 é referencial. De fato, a porcentagem de respostas N2 aumentou de 26% para um N2 não referencial no questionário anterior para 55% para um N2 referencial, corroborando, assim, a influência da informação da referencialidade na análise das orações relativas.

O princípio da Referencialidade foi investigado em português brasileiro primeiramente por Maia e Finger (2007). Os autores conduziram um questionário e investigaram a influência de vários tipos de sintagmas complexos no processamento de orações relativas, além de focalizarem a influência de fatores prosódicos nas preferências de aposição. No que tange ao sintagma nominal do tipo substância, os autores investigaram o processamento de sentenças como “O técnico fez críticas à antena de metal que oxida”, em que apenas o N1 é referencial, e não manipularam a referencialidade do N2. Os resultados seguiram a linha do que é proposto por Gilboy *et al.* (1995): diante de um N1 referencial e de um N2 referencial, falantes de português brasileiro preferiram a aposição da oração relativa ao N1.

Há evidências, portanto, para o princípio da Referencialidade, e elas remetem basicamente ao momento *off-line* do processamento. Uma questão que aparece, então, neste contexto, é se tal informação poderia influenciar a compreensão *on-line* da oração relativa, o que remete para uma reflexão mais ampla sobre a natureza dos processos de associação e interpretação assumidos pelo *Construal*. Veja-se que esta hipótese não propõe um mecanismo de análise específico para o processamento de sintagmas adjuntos, não postulando de forma clara como ocorre o processo de interpretação nem esclarecendo quão persistente seria a subespecificação sintática no decorrer do processamento. É uma questão aberta para o *Construal*, portanto, quão rápida seria a

interpretação responsável por conferir uma análise específica ao material subespecificado, o que decorre do fato de não se ter clareza quanto à forma pela qual o *parser* tem acesso à informação não estrutural que é responsável por determinar uma análise. No processamento de sintagmas primários, em específico, no processo de reanálise, o acesso à informação não estrutural pelo *parser* é viabilizado pelo processador temático (RAYNER; CARLSON; FRAZIER, 1983; FRAZIER, 1990), que traduz em um vocabulário legível para o *parser* informações concernentes ao discurso e à plausibilidade, por exemplo. Seria preciso também tornar explícito como o *parser* enxerga informações não sintáticas que são responsáveis por guiá-lo na determinação de análise de um sintagma secundário (associação > oposição).

Considerando, especificamente, o que poderia conduzir ou encorajar o *parser* a avançar de uma simples associação para a oposição de um sintagma, Swets, Desmet, Clifton e Ferreira (2008) enfatizam o papel da demanda da tarefa experimental, propondo um diálogo entre a Hipótese de *Construal* e a Teoria do *Good-Enough* (FERREIRA; BAYLEY; FERRARO, 2002; FERREIRA, 2003; FERREIRA; PATSON, 2007; KARIMI; FERREIRA, 2015). O fundamento desta teoria é o seguinte: o processamento da linguagem não necessariamente envolve a construção de representações linguísticas acuradas, detalhadas e completas, e isso seria uma decorrência do fato de que, em uma situação real de comunicação, o objetivo do sistema de compreensão não seria produzir representações ótimas a partir de uma computação sintática minuciosa, mas, na verdade, entregar uma interpretação que é apenas boa o bastante (*good-enough*) para permitir que o sistema de produção gere uma resposta apropriada, contribuindo para o sucesso da interação. Nessa linha de raciocínio, a profundidade de processamento do material linguístico (em termos do grau de comprometimento sintático) estaria relacionada ao objetivo da tarefa na qual os sujeitos estão engajados. O “*good-enough*” em situação real de comunicação pode deixar de sê-lo em uma situação de tarefa experimental, por exemplo.

A proposta de Swets *et al.* (2008) é a de que, posteriormente à associação inicial do sintagma secundário, a busca por informações que auxiliem a resolução da ambiguidade pode ser gradual, podendo se estender ou não até a determinação de uma análise de acordo com os objetivos do leitor. Nesse sentido, na ausência de um material desambiguador e de uma pressão, em termos experimentais, para uma

aposição específica, a consulta a várias fontes de informação poderia ser apenas superficial. Conforme os autores, a pressão para a realização de uma aposição poderia advir do tipo de questão de compreensão presente em uma tarefa experimental: dada uma ambiguidade quanto à aposição de uma oração relativa, tal como em “Alguém atirou no empregado da atriz que estava na varanda”, sujeitos em uma tarefa com questão de compreensão focalizando a oração relativa tendem a realizar uma aposição mais imediata, o que não ocorre no caso de uma tarefa com questão de compreensão superficial, isto é, que não direciona para a interpretação da oração relativa, na qual os sujeitos tendem, em verdade, a deixar uma análise subespecificada.

É importante considerar, no entanto, as críticas tecidas por Frazier (2008) à Teoria do *Good-Enough*. Conforme a autora, está claro que o objetivo do sistema de compreensão da linguagem não é computar uma representação sintática, mas esta computação seria um meio necessário para se alcançar a mensagem pretendida pelo falante. A sintaxe não seria apenas uma questão de “*niceties of high style*”, mas uma forma de individualizar e acessar entidades e eventos, sem a qual a interpretação poderia ser “*a salad of word meanings*”. Quanto à subordinação do processamento linguístico à demanda da tarefa experimental, a autora questiona a ideia de um controle estratégico da computação sintática e diz que poderia até mesmo não ser apropriado falar em “objetivo” do processador sintático, já que esse seria um sistema de inferência automático. A autora especula, porém, que talvez a fase mais interpretativa do processamento esteja mais aberta a efeitos de profundidade de processamento em dependência da tarefa experimental.

O debate está aceso e o presente estudo insere-se nessa discussão teórica geral na medida em que se propõe a investigar aspectos como: subespecificação inicial de análise e determinação de análise através de informação não estrutural no processamento de orações relativas em português brasileiro.

3 Experimentos

Os experimentos realizados por Gilboy *et al.* (1995) eram de medidas *off-line*, dando ênfase ao papel da referencialidade no processo de interpretação das orações relativas. Porém, para verificar até que ponto essa informação poderia influenciar o processamento *on-line* das

orações relativas, inserindo-se mais diretamente na discussão sobre o grau de subespecificação no processamento sintático que subjaz à noção de associação proposta pelo *Construal*, optou-se por investigar não só o momento mais reflexivo da compreensão, mas, sobretudo, o momento reflexo. Nesse sentido, apresentaremos a seguir dois experimentos *on-line* que conduzimos para compreender a natureza do fenômeno em questão.

3.1 Leitura automonitorada

Neste experimento, manipulamos a referencialidade do N2 e o gênero do particípio da oração relativa, originando as seguintes condições experimentais:

a) N2 Não Referencial / Aposição ao N1 (N2NR/N1)

O policial apreendeu/ a bolsa de couro/ que foi irregularmente importada/ pela empresa.

b) N2 Não Referencial / Aposição ao N2 (N2NR/N2)

O policial apreendeu/ a bolsa de couro/ que foi irregularmente importado/ pela empresa.

c) N2 Referencial / Aposição ao N1 (N2R/N1)

O policial apreendeu/ a bolsa do couro/ que foi irregularmente importada/ pela empresa.

d) N2 Referencial / Aposição ao N2 (N2R/N2)

O policial apreendeu/ a bolsa do couro/ que foi irregularmente importado/ pela empresa.

Tomamos como variável dependente o tempo de leitura do segmento três, que contém o material desambiguador, e partimos da hipótese geral de uma influência *on-line* da referencialidade do N2 e da informação de gênero, assumindo, então, uma preferência geral pela aposição alta da oração relativa quando o N2 é não-referencial mas uma atenuação dessa preferência quando o N2 é referencial, com o gênero do particípio sinalizando essas preferências. Nessa linha, elencamos as seguintes previsões:

- 1) tempos de leitura mais lentos em (b) do que em (a);
- 2) tempos de leitura mais rápidos em (d) do que em (c), pois, em (d), as duas variáveis convergem para a aposição local;
- 3) tempos de leitura mais rápidos em (d) do que em (b), pois, em (b), o N2 não referencial não favorece a aposição local forçada pelo gênero;
- 4) finalmente, tempos de leitura mais lentos em (c) do que em (a), pois (c) contém um N2 referencial, que, potencialmente, favorece uma aposição local, mas o gênero guia o *parser* para a aposição não local.

3.1.1 Método

a) Participantes

Participaram voluntariamente do experimento trinta e dois graduandos de universidades estaduais e particulares de Campina Grande (PB) (dos cursos de Direito, Enfermagem e Engenharia Química), todos falantes nativos de Português Brasileiro e com idade média de 21 anos.

b) Material

O material compreendeu quatro conjuntos experimentais, cada um com dezesseis sentenças experimentais e trinta e duas sentenças distratoras: atendendo ao *design* 2x2, foram geradas quatro condições experimentais, exemplificadas anteriormente, e, em cada conjunto experimental, foram inseridas quatro sentenças experimentais por condição e o dobro de sentenças distratoras. O modo de distribuição dos itens experimentais em quatro conjuntos respondeu à lógica do *design* do tipo quadrado latino e do *design* do tipo intrassujeito, de forma que todos os participantes entraram em contato com todas as condições experimentais, mas não com mais de uma versão de um mesmo item experimental.

No que se refere à estrutura e ao modo de segmentação das sentenças experimentais, elas apresentavam a seguinte estruturação: SN + V / SN complexo (N1 de N2) / oração relativa / Agente da passiva. A estrutura do segmento 2 modificava-se entre as condições: nas condições

N2NR, o N2 não era antecedido por um Determinante, diferentemente do que era observado nas condições N2R. Ainda no que tange ao SN complexo, o N1 e o N2 sempre diferiam quanto ao gênero: quando o N1 correspondia ao gênero feminino, o N2 correspondia ao gênero masculino, e vice-versa. A estrutura do segmento 3, que é o segmento crítico, também se alterava entre as condições: nas condições N1, o gênero do particípio concordava com o N1, direcionando para a aposição não local da oração relativa, e nas condições N2, o gênero do particípio concordava com o N2, direcionando para a aposição local da oração relativa. O segmento 4 correspondia ao segmento pós-crítico e, também, ao segmento final da sentença. Após este segmento, aparecia, em todos os itens experimentais, uma questão de compreensão, que focalizava a aposição da oração relativa ao N1, tal como “A bolsa foi importada?”, para as sentenças mostradas anteriormente. Como a questão focalizava o N1, nas condições (a) e (c), a resposta correta era SIM, e, nas condições (b) e (d), a resposta correta era NÃO.

c) Procedimentos

O experimento foi realizado por meio da técnica on-line de leitura automonitorada na modalidade não cumulativa. Os participantes liam sentenças exibidas de maneira segmentada na tela do computador e apertavam a tecla L para passar de um segmento a outro, respondendo SIM ou NÃO à questão de compreensão que aparecia depois de cada sentença. O aparato experimental abarcou um *Macbook Apple*, cujo sistema operacional suporta o programa *Psyscope*, o qual foi utilizado para programar e rodar o presente experimento.

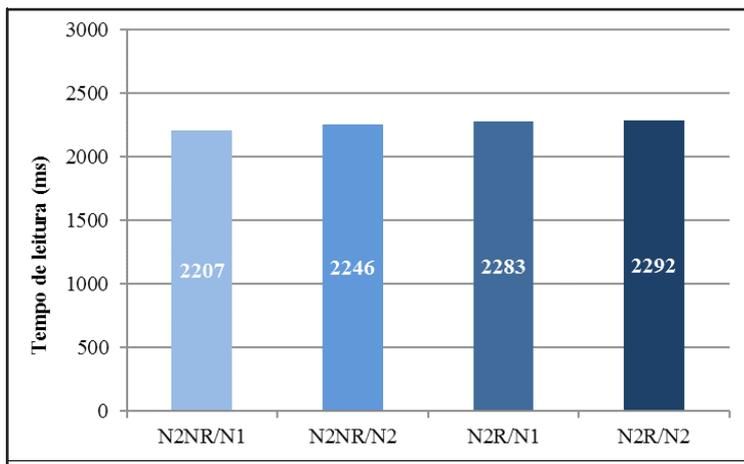
Nas sessões experimentais, os participantes eram testados individualmente, em uma sala isolada, e recebiam, antes da realização da tarefa, instruções da experimentadora no que tange ao mecanismo geral da tarefa. Além dessa instrução, os sujeitos participavam de uma etapa de prática, na qual entravam em contato com oito sentenças que possuíam estruturas distintas das estruturas das sentenças experimentais. Cada sessão experimental durou cerca de 20 minutos e os sujeitos não reportaram dificuldades na execução da tarefa.

3.1.2 Resultados

Os resultados obtidos na análise estatística do segmento crítico não foram coerentes com as previsões: a análise da variância (ANOVA)

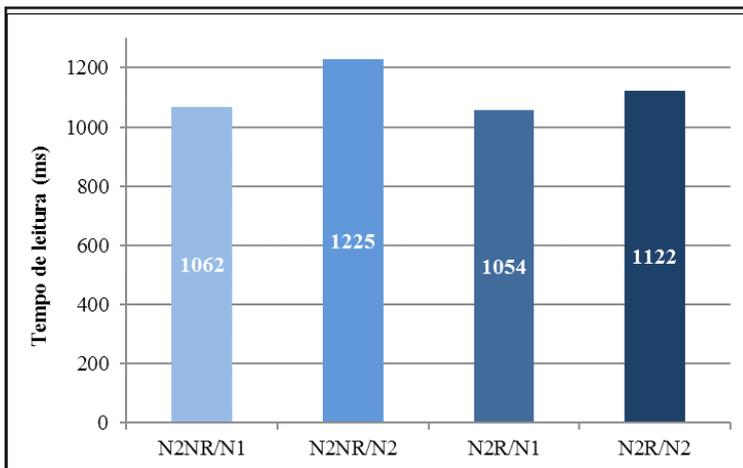
não revelou efeito principal da referencialidade do N2 ($F(1,31)=0,459$; $p < 0,5$) nem de gênero ($F(1,31)=0,076$; $p < 0,7$), e também não revelou efeito de interação entre as variáveis ($F(1,31)=0,019$; $p < 0,8$). Os tempos médios de leitura podem ser vistos no gráfico 1:

GRÁFICO 1 – Tempos médios de leitura do segmento crítico



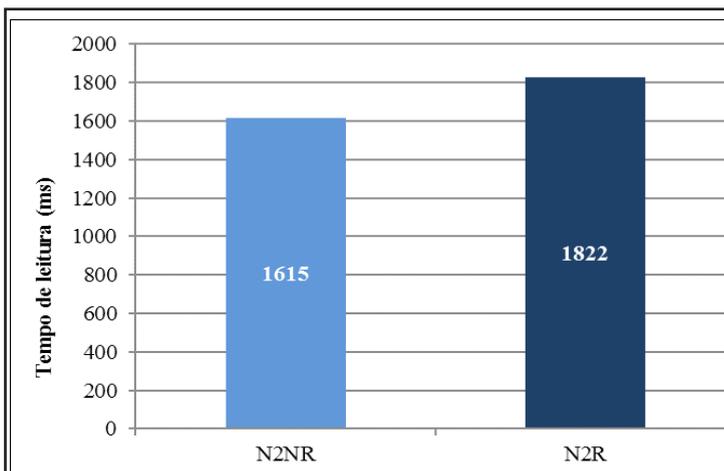
Em busca de um efeito *spillover*, analisamos os tempos de leitura do segmento pós-crítico. A ANOVA revelou efeito principal de gênero ($F(1,31)=8,34$; $p < 0,007$), evidenciando maiores tempos de leitura nas condições N2 (gênero forçando aposição ao N2), mas não revelou efeito principal da referencialidade do N2 ($F(1,31)=1,29$; $p < 0,2$) nem efeito de interação entre as variáveis ($F(1,31)=1,51$; $p < 0,2$). Considerando o efeito de gênero, as análises com Teste T revelaram uma diferença significativa apenas na comparação N2NR/N1 vs N2NR/N2 ($t(31)=2,79$; $p < 0,008$), mas não na comparação N2R/N1 vs N2R/N2 ($t(31)=1,28$; $p < 0,21$). Os tempos médios de leitura podem ser vistos no gráfico 2:

GRÁFICO 2 – Tempos médios de leitura do segmento pós-crítico / final



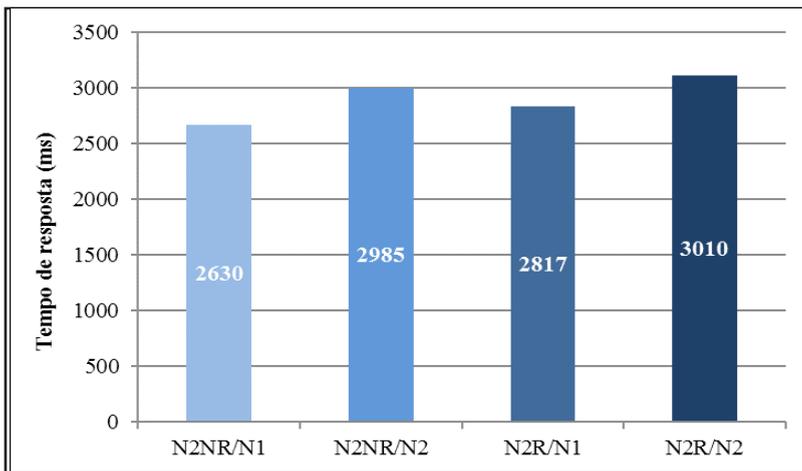
Buscando entender a ausência do efeito de referencialidade do N2, verificamos, também, o comportamento dos sujeitos diante do SN complexo em si. Comparando os tempos de leitura do SN complexo na condição N2NR com os tempos de leitura do SN complexo na condição N2R, o Teste T revelou que os tempos são significativamente maiores na condição N2R ($t(31)=-2,41$; $p<0,02$). Esse resultado evidencia que os sujeitos computaram o determinante na condição N2R, embora ele não tenha afetado o processamento da oração relativa subsequente. Os tempos médios de leitura podem ser vistos no gráfico 3:

GRÁFICO 3 – Tempos médios de leitura do SN complexo



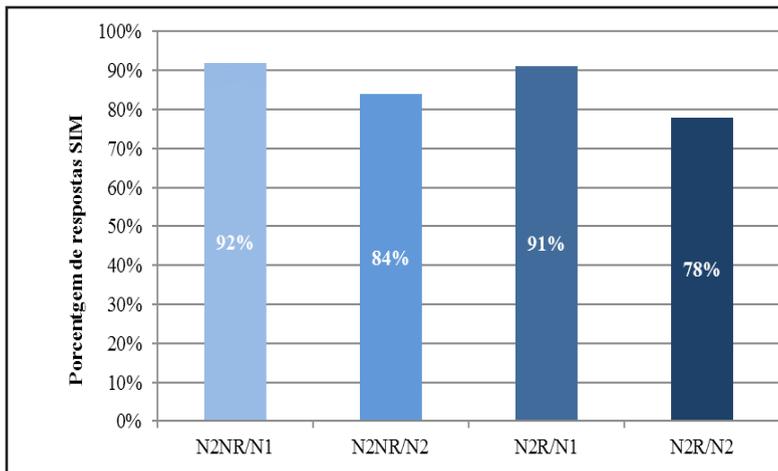
Diante da ausência de efeito da referencialidade do N2 nestes dados *on-line*, analisamos os tempos de respostas e as respostas à questão de compreensão. Vendo, primeiramente, os tempos de resposta, a ANOVA não revelou efeito principal da referencialidade do N2 ($F(1,31)=0,570$; $p<0,4$), nem de efeito de interação entre as variáveis ($F(1,31)=0,467$; $p<0,4$), mas revelou um efeito marginal da informação de gênero ($F(1,31)=3,65$; $p<0,06$), com efeito marginal emergindo na comparação N2NR/N1 vs N2NR/N2 ($t(31)=1,79$; $p<0,08$), mas sem efeito significativo ou marginal na comparação N2R/N1 vs N2R/N2 ($t(31)=1,12$; $p<0,27$). Nessa linha, embora de forma não significativa em termos estatísticos, os tempos para responder à questão foram, de modo geral, maiores nas condições N2, como pode ser visto no gráfico 4:

GRÁFICO 4 – Tempos médios de resposta à questão de compreensão



Vejam os padrões de respostas. Através de testes Qui-quadrado, encontrou-se um número significativamente maior de respostas SIM do que de respostas NÃO em todas as condições experimentais (N2NR/N1: $X^2=91,125$; $p<0,05$; N2NR/N2: $X^2=60,5$; $p<0,05$; N2R/N1: $X^2=87,78$; $p<0,05$; N2R/N2: $X^2=40,5$; $p<0,05$). Visualizou-se, porém, um efeito principal de gênero ($X^2=390,94$, $p<0,0003$), com o gênero forçando a oposição ao N2 aumentando significativamente a probabilidade de respostas N2 ($\beta=0,7817$; $SE=0,4096$; $z=1,909$; $p<0,05$). As porcentagens de respostas SIM (N1) podem ser vistas no gráfico 5:

GRÁFICO 5 – Percentagem de respostas SIM à questão de compreensão



Discutiremos esses resultados no subtópico a seguir.

3.1.3 Discussão

Com base nos dados obtidos, vemos que apenas a previsão (1) foi corroborada. Verificamos que, em estruturas com SNs complexos do tipo “substância”, não há uma preferência sistemática pela aposição local da oração relativa, em oposição ao que seria predito pelo princípio da Aposição Local (FRAZIER, 1979), mas há uma preferência geral pela aposição não local, já encontrada anteriormente por Gilboy *et al.* (1995) e Maia e Finger (2007) em uma etapa mais interpretativa da compreensão. Supõe-se que a base dessa preferência de análise tem relação com a natureza semântica do SN complexo “substância”. Em SNs como “a bolsa de (do) couro”, vemos que a relação entre o N1 e o N2, mediada pela preposição, é a de que “X é feito da substância Y”, e essa relação não favorece (ou mesmo bloqueia) a introdução de duas entidades discursivas no modelo mental, favorecendo, na verdade, uma análise em que o N2 funciona apenas como modificador do N1, ambos correspondendo a uma única entidade discursiva. Os dados sugerem que o *parser* tem um acesso rápido à restrição imposta por essa relação

semântica, daí a resistência¹ que os sujeitos apresentam à construção com N2 referencial.

Tal resistência e a ausência de efeito da referencialidade do N2 indicam que a implicação discursiva da adição do determinante ao N2 (a sua introdução no modelo discursivo como entidade independente do N1 mas a ele relacionado via modificação) não foi a análise imediatamente vista pelo *parser*. De fato, considerar o N2 como uma entidade discursiva parece requerer um processo mais reflexivo, dado que o artigo definido em si carrega pressuposição (HAWKINS, 1978; HEIM, 1982) e o experimento não contou com contexto prévio. Uma objeção possível seria a de que o N1 também é definido e também está sendo introduzido no modelo discursivo mas, ainda assim, ele é selecionado como local de aposição. Porém, é preciso observar que, no SN complexo em foco, o N1 e o N2 exibem *status* sintático e semântico distintos. Em termos gerais, o N1 é o núcleo do SN complexo e é alvo de modificação restritiva, enquanto o N2 justifica sua presença apenas em termos do N1. A nossa hipótese é a de que essa “subordinação” sintático-semântica do N2 (em relação ao N1) pode dificultar o seu processo de acomodação (HEIM, 1982), isto é, pode dificultar a sua introdução no modelo mental em construção para a sentença. O *parser*, portanto, enxergaria tal subordinação, mas não a dimensão discursiva do N2 referencial.

Reconhecendo essa preferência geral pela aposição não local, devemos refletir, agora, sobre o momento em que ela aparece no curso da compreensão. Para tanto, podemos olhar para o efeito de gênero. Em experimentos psicolinguísticos, o uso de material desambiguador é elucidativo no sentido de poder indicar uma preferência de análise: se uma análise X foi selecionada em um ponto de ambiguidade, o material desambiguador, quando contrário à análise X inicial, pode gerar um efeito *garden-path* e um processo de reanálise, o que causa um maior custo computacional em comparação a um caso em que o material desambiguador apenas corrobora a análise X. No nosso caso, o gênero

¹ Nós estamos interpretando os maiores tempos de leitura na condição com N2 referencial como uma resistência pois não se observou um efeito dessa referencialidade no processamento da oração relativa em si. Caso essa influência tivesse sido observada, uma interpretação alternativa seria a de que os tempos estão refletindo um custo associado à inserção e à integração de uma nova entidade no modelo discurso (cf. Warren; Gibson, 2002; Grodner; Gibson; Watson, 2005).

é o responsável pela sinalização de que a análise X, aqui, é aposição ao N1. O efeito de gênero foi encontrado no segmento pós-crítico, que correspondia também ao segmento final, e poderia ser interpretado ou como um efeito *spillover* de um processamento ocorrido no segmento crítico em si ou como um efeito tardio no sentido de ter sido considerado apenas em uma fase mais final do processamento, talvez como parte de um *wrap-up process* (JUST; CARPENTER, 1980), que envolve a integração de várias fontes de informação, chegando-se à compreensão final da sentença. Essa discussão se mostra ainda mais pertinente ao se considerar que já existe, na literatura, um debate quanto à sensibilidade do *parser* a essa informação de gênero, com estudos propondo que ela poderia ser tardiamente acessada (DE VINCENZI; DI DOMENICO, 1999) ou mesmo desconsiderada (BRYSSBAERT; MITCHELL, 2000) no processamento. Nesse sentido, o efeito de gênero, por si só, não nos permite dizer, com categoria, se a preferência de aposição ao N1 foi imediata ou se houve uma subespecificação inicial de análise mais prolongada.

Mas se assumirmos que houve uma preferência mais imediata pelo N1, o efeito de gênero poderia ser entendido pelo menos de duas formas, dada a falta de acurácia no padrão de respostas das condições N2 (gênero forçando a aposição ao N2): ele poderia ser o resultado de uma resistência em considerar a possibilidade de apor a oração relativa ao N2 ou poderia ser visto como um sinal de reanálise, mas, nesse caso, de uma reanálise parcial, pois, ao final, é o N1 que acaba sendo majoritariamente escolhido. No âmbito da Teoria do *Good-Enough* (CHRISTIANSON; HOLLINGWORTH; HALLIWELL; FERREIRA, 2001; RIBEIRO, 2008, 2010, 2012), tem sido considerada a ideia de má interpretação de sentenças *garden-path* associada com realização de reanálises incompletas, assumindo que o princípio da Composicionalidade, segundo o qual o significado de uma sentença resulta da soma dos significados de suas partes, poderia ser violado no processamento.² Assim, os sujeitos poderiam não ter nem iniciado a reanálise ou tê-la iniciado mas não

² A Teoria do *Garden-Path* (FRAZIER, 1979) não assume esta hipótese, defendendo, na verdade, uma construção composicional do significado das sentenças. Nessa linha, após um efeito *garden-path*, para que a sentença seja devidamente interpretada, é necessário que haja um processo de reanálise pleno.

finalizado, tendo sido o gênero, de qualquer forma, desconsiderado na fase mais final da compreensão, como evidenciado pelas respostas.

Para tentar esclarecer o efeito de gênero e compreender melhor a influência da referencialidade marcada no N2, conduzimos um experimento com rastreamento ocular. Essa técnica é, em geral, caracterizada como mais natural (pela apresentação não segmentada das sentenças, por exemplo) e mais sensível ao processamento sintático inicial em comparação com outras técnicas, como a leitura automonitorada. Acreditamos, portanto, que, por sua própria natureza, o rastreamento ocular possa nos ajudar a esclarecer os dados obtidos neste experimento.

3.2 Rastreamento ocular

Neste segundo experimento, manipulamos, também, a referencialidade do N2 e o gênero do participio da oração relativa. As condições experimentais foram as mesmas do teste anterior:

a) N2 Não Referencial / Aposição ao N1 (N2NR/N1)

O policial confiscou a bolsa de couro que foi importada irregularmente pela empresa.

b) N2 Não Referencial / Aposição ao N2 (N2NR/N2)

O policial confiscou a bolsa de couro que foi importado irregularmente pela empresa.

c) N2 Referencial / Aposição ao N1 (N2R/N1)

O policial confiscou a bolsa do couro que foi importada irregularmente pela empresa.

d) N2 Referencial / Aposição ao N2 (N2R/N2)

O policial confiscou a bolsa do couro que foi importado irregularmente pela empresa.

Consideramos como região de interesse a forma participial do verbo presente na oração relativa. Como variável dependente, consideramos, inicialmente, o tempo de primeira leitura, que, sendo associada às etapas mais iniciais do processamento sintático, corresponde à soma da duração das primeiras fixações em uma região desde a primeira fixação nessa região até ao seu abandono para a esquerda ou para a direita.

Partindo da hipótese de um efeito imediato da referencialidade do N2 e do gênero do participio da oração relativa, fizemos as seguintes previsões:

- 1) Tempos mais lentos em (b) do que em (a);
- 2) Tempos mais lentos em (c) do que em (a);
- 3) Tempos mais lentos em (c) do que em (d);
- 4) Tempos mais rápidos em (d) do que em (b).

Caso o efeito da referencialidade do N2 e o efeito do gênero não se mostrassem de imediato, poderíamos encontrar as diferenças previstas acima em medidas mais tardias, como no tempo de segunda leitura (que abarca os tempos relativos a todas as releituras da região, incluindo a segunda, a terceira ou e as demais vezes que o leitor entrou nessa região) e no tempo de leitura total (que corresponde à soma dos valores da primeira e da segunda leituras).

3.2.1 Método

a) Participantes

Participaram voluntariamente do experimento vinte e oito graduandos da Universidade Federal do Rio de Janeiro, todos falantes nativos de português brasileiro e com idade média de 22 anos. Os sujeitos receberam duas horas em sua carga horária do Curso de Letras pela participação voluntária no teste.

b) Material

Assim como no experimento anterior, o material abarcou quatro conjuntos experimentais, cada um com dezesseis sentenças experimentais (quatro por condição) e trinta e duas sentenças distratoras. O modo de distribuição dos itens experimentais respondeu à lógica do *design* do tipo quadrado latino e do *design* do tipo intrassujeito, de forma que todos os participantes entraram em contato com todas as condições experimentais, mas não com mais de uma versão de um mesmo item experimental.

Os itens experimentais tinham a mesma estrutura das sentenças do experimento anterior, com apenas uma modificação: o advérbio foi

colocado, agora, após o material desambiguador.³ Com relação aos itens lexicais, em todos eles, o N1 correspondia ao gênero feminino e o N2 correspondia ao gênero masculino. Após cada sentença experimental, aparecia uma questão de compreensão (focalizando a oração relativa) e duas opções de resposta: N1 e N2, tal como “O que a empresa importou? () Bolsa () Couro”,⁴ para os exemplos apresentados. Nas sentenças experimentais, as respostas eram apresentadas sempre com o N1 à esquerda e o N2 à direita; já nas sentenças distratoras, metade das respostas corretas aparecia à esquerda e a outra metade aparecia à direita.

c) Procedimentos

O aparelho utilizado na coleta dos dados foi um rastreador modelo *EyeLink* 1000, com uma câmera com acurácia de 1000Hz, configurada para gravação monocular, acoplada a uma tela de 32 polegadas, com resolução de 1920x1080 px. Foi utilizado aparato para a cabeça e a testa do voluntário, enquanto seus movimentos oculares foram gravados durante a leitura das frases. O programa usado para apresentação e registro dos estímulos foi o “*EyeTracker* 0.7.10m” e a análise foi feita utilizando-se os programas “*Robodoc.py*” e “*EyeDry*”.

O experimento foi aplicado na sala do Laboratório de Psicolinguística Experimental – Lapex, na Faculdade de Letras da UFRJ. Cada participante era instruído a ler frases, enquanto seus olhos eram monitorados pelo aparelho. Após a instrução da tarefa, o participante era posicionado no equipamento. O experimentador ajustava a cabeça e a

³ Intuitivamente, achamos que, com a antecipação do material desambiguador, o efeito do gênero poderia ser mais rápido e mais forte. Ferreira e Henderson (1991) propõem, nessa linha, que a distância entre o ponto em que uma análise é atribuída e o ponto em que o material desambiguador aparece pode influenciar a força da análise inicial na memória e, conseqüentemente, o processo de reanálise: quanto maior a distância, maior a força da primeira análise na memória, o que não favorece a reanálise. Em um estudo de leitura automonitorada (não publicado), fizemos essa alteração, mas nenhuma diferença significativa quanto ao efeito do gênero foi observada.

⁴ Respondendo a um dos revisores anônimos: usamos um tipo de questão diferente do utilizado no experimento anterior pois, em um experimento ainda não publicado, encontramos evidências de que o tipo de questão pode influenciar o comportamento dos sujeitos, com uma questão como a do experimento 1, exemplo, diminuindo a probabilidade de respostas N2.

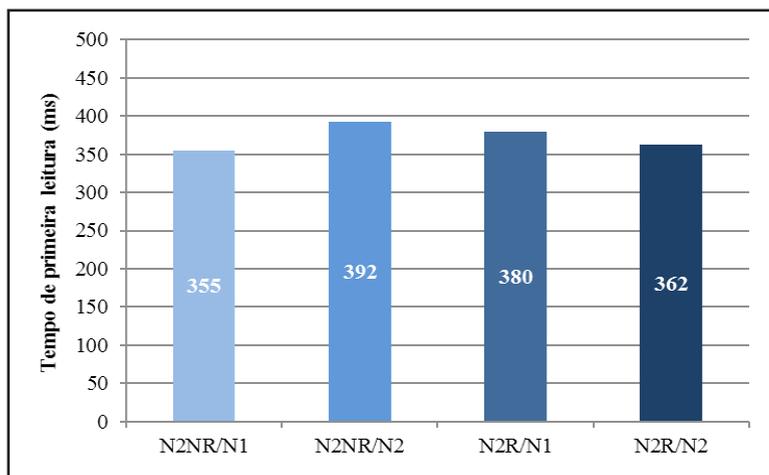
testa do voluntário no aparato, pois era importante que o voluntário ficasse com a cabeça imóvel durante a aplicação do teste. Após o devido ajuste, o voluntário tinha a sua pupila calibrada e validada. Nesse processo, o participante era instruído a olhar fixamente para um círculo preto que aparecia inicialmente no centro da tela e movia-se para 9 pontos diferentes. Em seguida, o participante passava pela fase de validação, na qual o aparelho checava todas as fixações oculares capturadas durante a calibragem. Essa fase era de extrema relevância, pois apenas depois da perfeita calibragem e validação da pupila do voluntário era possível continuar com o procedimento.

Antes de cada frase, no canto medial esquerdo da tela, aparecia um quadrado preto indicando que o participante deveria olhar fixamente para este local, para que a frase pudesse surgir na tela. A frase só aparecia se o participante estivesse fixando o olhar no quadrado preto. Isso acontecia sempre antes de cada frase ser apresentada. Após a leitura da frase, o participante era instruído a apertar um botão no *joystick* (controle), autorizando o aparecimento da questão interpretativa. O participante, então, respondia apertando o botão esquerdo ou direito, e em seguida aparecia um novo quadrado preto, indicando o aparecimento de uma nova frase, e assim por diante. Cada participante realizava a leitura de 96 frases (16 frases experimentais e 32 frases distratoras, com suas respectivas questões interpretativas). Era necessário que o experimentador permanecesse na sala juntamente com o participante, para manuseio do monitor principal, através do qual eram realizados o processo de calibragem e validação, além do aceite das fixações dos voluntários. No processo de aceite das fixações oculares, o experimentador clicava com o mouse sobre a pupila do voluntário, sempre que estava certo de que o participante estava, de fato, olhando para a área correta. Tal procedimento acontecia sempre antes de o quadrado preto surgir e consistia no aparecimento de um círculo preto no centro da tela, onde o participante tinha que olhar fixamente, para que o experimentador pudesse aceitar sua fixação. A duração de cada teste era em torno de 30 minutos. Antes de cada experimento era aplicada uma prática, com o objetivo de treinamento e apresentação da tarefa, com duração de 10 minutos.

3.2.2 Resultados

Focalizando os dados de primeira leitura, a ANOVA⁵ revelou um efeito principal de gênero ($F(1,99)= 8,910; p<0,003$), um efeito principal da referencialidade do N2 ($F(1,99)= 7569,304; p<0,001$) e um efeito de interação das variáveis ($F(1,99)=29,024; p<0,001$). A análise estatística por pares com Teste T evidenciou como significativas as seguintes comparações: N2R/N1 vs N2R/N2 ($t(99)=28,899; p<0,001$), com a condição N2R/N1 exibindo tempos de leitura maiores; N2NR/N1 vs N2R/N1 ($t(99)=-59,700; p<0,001$), com a condição N2R/N1 exibindo tempos de leitura maiores; e N2NR/N2 vs N2R/N2 ($t(99)=64,116; p<0,001$), com a condição N2NR/N2 exibindo tempos de leitura maiores. A comparação N2NR/N1 vs N2NR/N2 não revelou uma diferença significativa ($t(99)=-1,291; p<0,1$), havendo, porém, uma diferença numérica na direção esperada. Estes dados mostram a pertinência das previsões (2), (3) e (4) que elencamos anteriormente. Os tempos médios de leitura em cada condição podem ser vistos no gráfico 6:

GRÁFICO 6 – Tempos médios de primeira leitura de “importado(a)”

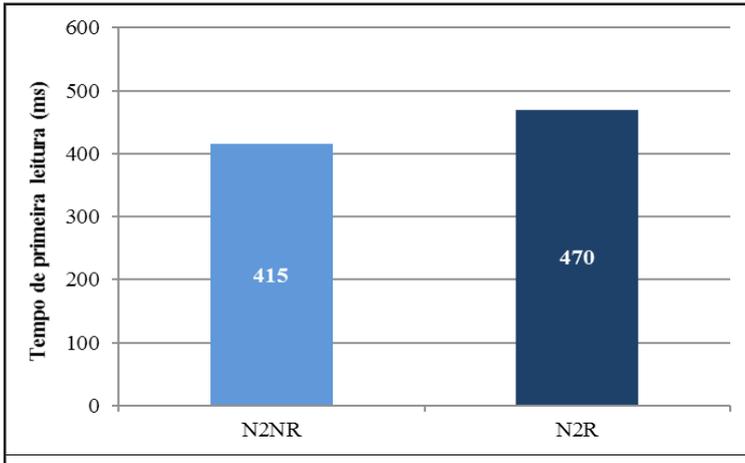


⁵ Os dados de todas as medidas do rastreamento ocular que consideramos para análise não apresentavam distribuição normal originalmente. Desta forma, para conduzir a ANOVA e o Teste T, procedemos a uma transformação dos dados, o que fizemos por meio do recurso “Transformação Box-Cox” disponibilizado pelo programa estatístico *Action*.

Analisando os tempos de segunda leitura, a ANOVA revelou um efeito principal de gênero ($F(1,67)=439,432$; $p<0,001$), um efeito principal da referencialidade do N2 ($F(1,67)=433,157$; $p<0,001$) e um efeito de interação das variáveis ($F(1,67)=422,247$; $p<0,001$). A análise estatística por pares com Teste T evidenciou como significativas as seguintes comparações: N2NR/N1 vs N2NR/N2 ($t(67)=21,138$; $p<0,001$), com a condição N2NR/N2 exibindo tempos de leitura maiores; N2R/N1 vs N2R/N2 ($t(67)=1,926$; $p<0,001$), com a condição N2R/N1 exibindo tempos de leitura maiores; e N2NR/N1 vs N2R/N1 ($t(67)=20,685$; $p<0,001$), com a condição N2R/N1 exibindo tempos de leitura maiores. A análise dos tempos de leitura total evidenciou efeito significativo de gênero ($F(1,99)=4,086$; $p<0,04$) e efeito de interação das variáveis ($F(1,99)=4,857$; $p<0,02$), com o Teste T revelando como significativas as seguintes comparações: N2NR/N1 vs N2NR/N2 ($t(99)=-3,071$; $p<0,002$), com a condição N2NR/N2 exibindo tempos de leitura maiores; e N2NR/N1 vs N2R/N1 ($t(99)=-2,126$; $p<0,03$), com a condição N2R/N1 exibindo tempos de leitura maiores.

Para entender melhor como os sujeitos lidaram com o N2 referencial em si, decidimos considerar uma segunda região de interesse: o sintagma preposicional (SP) do SN complexo – “de (do) couro”, por exemplo. Analisando os tempos de primeira leitura através do Teste T, visualizamos que as condições com N2 referencial geram um custo de processamento maior do que as condições com N2 não referencial ($t(193)=-2,140$; $p<0,03$), como podemos observar no gráfico 7:

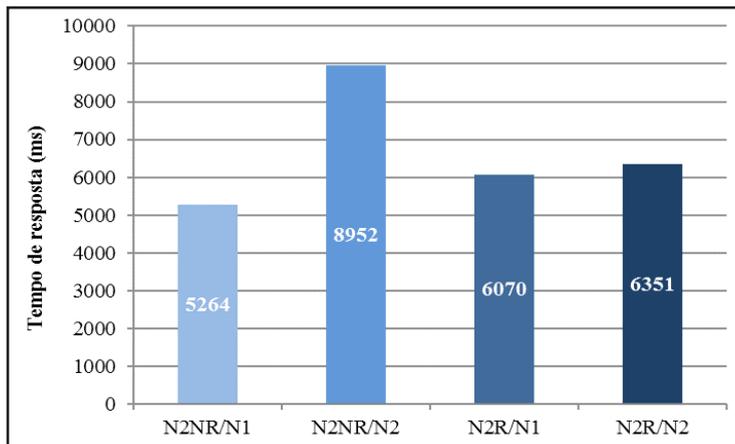
GRÁFICO 7 – Tempos médios de primeira leitura do SP



Esse mesmo custo adicional visualizado nas condições N2R apareceu nos tempos de leitura total do SP ($t(193)=-3,815$, $p<0,0001$), mas não nos tempos de segunda leitura ($t(145)=-1,323$; $p<0,1$).

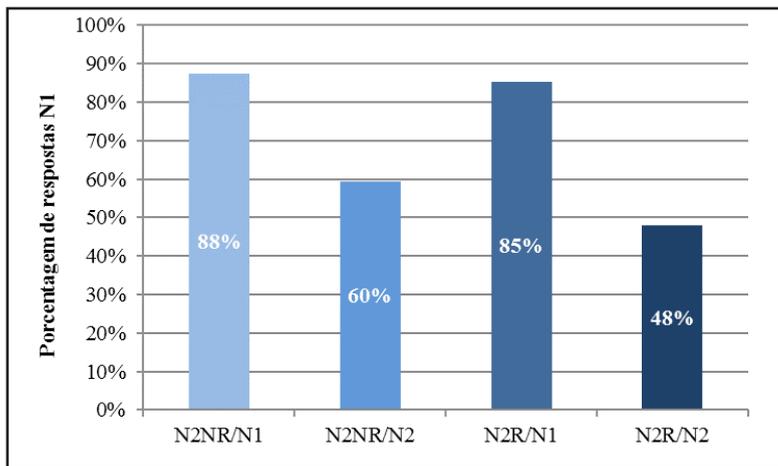
Finalmente, analisamos, também, os dados das respostas. Quanto aos tempos de resposta, a ANOVA revelou efeito significativo de gênero ($F(1,111)=3428,890$; $p<0,001$), efeito significativo da referencialidade do N2 ($F(1,111)=3471,358$; $p<0,001$) e efeito de interação das variáveis ($F(1,111)=2069,417$; $p<0,001$). A análise estatística por pares efetuada com Teste T revelou efeitos significativos nas comparações relevantes: N2NR/N1 vs N2NR/N2 ($t(111)=152,373$; $p<0,001$), com a condição N2NR/N2 exibindo tempos mais lentos; N2R/N1 vs N2R/N2 ($t(111)=52,190$; $p<0,001$), com a condição N2R/N2 exibindo tempos mais lentos; N2NR/N1 vs N2R/N1 ($t(111)=-52,716$; $p<0,001$), com a condição N2R/N1 exibindo tempos mais lentos; e N2NR/N2 vs N2R/N2 ($t(111)=-175,743$; $p<0,001$), com a condição N2NR/N2 exibindo tempos mais lentos. Os tempos médios de resposta podem ser vistos no gráfico 8:

GRÁFICO 8 – Tempos de resposta à questão de compreensão



Quanto às respostas em si, testes de Qui-quadrado de aderência revelaram um maior número de respostas N1 nas condições N2NR/N1 ($X^2=63$; $p<0,001$), N2NR/N2 ($X^2=3,9729$; $p<0,04$) e N2R/N1 ($X^2=54,32$; $p<0,001$), não havendo diferença, porém, entre o número de respostas N1 e N2 na condição N2R/N2 ($X^2=0,1428$; $p<0,7$). Encontramos, nesta linha, um efeito significativo de gênero ($X^2=485,01$; $p<0,001$), com a probabilidade de obter uma resposta N2 sendo significativamente maior quando o gênero força a aposição ao N2 ($\beta=1,5629$; $SE=0,3450$; $z=4,531$; $p<0,001$), como podemos observar no gráfico 9:

GRÁFICO 9 – Porcentagem de respostas N1 à questão de compreensão



Esse conjunto de dados será discutido no subtópico a seguir.

3.2.3 Discussão

Os resultados deste experimento revelaram um efeito imediato do gênero do particípio e da referencialidade do N2: quando o N2 é referencial, há uma preferência imediata pela aposição da oração relativa ao N2, com a condição com gênero forçando a aposição ao N1 causando um maior custo no processamento. Já quando o N2 não é referencial, sugerimos que haveria uma preferência imediata pela aposição ao N1. Nesta linha, a associação inicial proposta pelo *Construal* (FRAZIER; CLIFTON, 1996) seria rapidamente resolvida, produzindo uma aposição. Esses dados são coerentes com o Princípio da Referencialidade (GILBOY *et al.*, 1995), segundo o qual, em caso de SNs complexos do tipo “substância”, a chance de aposição local da oração relativa seria muito maior quando o N2 é referencial do que quando o N2 não é referencial. Considerando a natureza “substância” do SN complexo que estamos focalizando, é interessante refletir sobre o custo adicional encontrado no SN complexo que contém um N2 referencial. No experimento anterior, este custo foi associado a uma resistência a um N2 referencial; porém, nesse segundo experimento, a referencialidade do N2 parece ter influenciado o processamento da oração relativa em si, de modo

que tal custo poderia ser associado a uma resistência e a uma tentativa de integração do N2 referencial no modelo discursivo. Neste sentido, o custo maior na condição N2R/N1 na região do participio (encontrado nos tempos de primeira leitura, nos tempos de segunda leitura e nos tempos de leitura total) estaria relacionado justamente à quebra da análise discursiva do N2: considerando que uma sentença como “O policial confiscou a bolsa do couro” não é bem formada em português brasileiro, um N2 referencial nesse tipo de SN complexo só é possível na presença de um modificador restritivo, que seria o responsável por estabelecer o seu referente (cf. Hawkins (1978) para uma discussão sobre casos de “*Referent-Establishing relative clauses*”, isto é, casos em que o SN “pede” uma oração relativa pois ele, sozinho, não consegue introduzir um referente), assim, quando o gênero do participio bloqueia esta análise, a integração discursiva do N2 fica prejudicada, com ele podendo ser reanalisado como não referencial (o que explica o número equiparável de respostas N1 nas condições N2NR/N1 e N2R/N1).

Considerando a interpretação acima, precisamos refletir mais sobre dois resultados: o maior custo encontrado na condição N2NR/N2 (sobretudo nos tempos de segunda leitura e nos tempos de leitura total) e o menor custo encontrado na condição N2R/N2 (sobretudo nos tempos de primeira leitura e de segunda leitura). No primeiro caso, a questão é se o custo pode ou não ser associado a um processo de reanálise, já que os dados das respostas revelaram ainda uma preferência pelo N1. É possível que, diante do N2 não referencial, os sujeitos tenham resistido a uma aposição local, realizando o processo de reanálise em alguns itens experimentais, mas em outros não; parece-nos legítimo considerar, também, a possibilidade de o gênero forçando a aposição ao N2 ter instaurado uma “dúvida” no processamento, com os maiores tempos de resposta nesta condição podendo estar associados a uma decisão de análise. Com relação ao menor custo na condição N2R/N2, que estaria relacionado à convergência das variáveis manipuladas, é curioso observar que, apesar da influência *on-line* da referencialidade do N2, houve um número equiparável de respostas N1 e N2 à questão de compreensão, já que esperávamos um maior número de respostas N2. Uma hipótese, aqui, seria a de que a oração relativa foi, de fato, aposta ao N2 nesta condição, com as respostas evidenciando que esta aposição não elimina a unidade conceitual correspondente ao SN complexo: o processo “a bolsa feita de (do) couro que foi importado irregularmente” implica a unidade-produto

“a bolsa de couro”, e, sendo o couro importado, a bolsa também o teria sido, com este processo inferencial podendo estar associado aos tempos de resposta mais lentos nesta condição.

Os resultados deste experimento são positivos pois evidenciam uma influência imediata dos dois fatores que focalizamos em nossa pesquisa: o gênero do particípio, que tem o potencial de nos revelar uma preferência de análise, e a referencialidade do N2, que nos permite investigar a influência de princípios não estruturais no processamento sentencial. Vimos que o processamento da oração relativa é influenciado pelo *status* referencial dos nomes que compõem o SN complexo, como predito pelo *Construal* e pelo Princípio da Referencialidade (GILBOY *et al.*, 1995; FRAZIER; CLIFTON, 1996), e que o processo de subespecificação envolvido no processamento de sintagmas secundários pode ser rapidamente resolvido quando há uma informação útil disponível (no nosso caso, a informação referencial já estaria disponível no SN complexo em si).

4 Discussão geral

Uma questão principal motivou a realização do experimento de leitura automonitorada: a informação não estrutural de referencialidade pode influenciar as decisões iniciais do *parser*? Os nossos resultados evidenciaram uma preferência geral pela aposição ao N1, com a referencialidade do N2, em particular, não exercendo uma influência na análise da oração relativa. Esses dados não nos permitem dizer, simplesmente, que a referencialidade não influencia o processamento inicial, haja vista que a preferência imediata pela aposição não local observada também residiria na própria natureza referencial do N1, como defendem Gilboy *et al.* (1995) para dados *off-line*, não sendo o fruto da aplicação sistemática de um princípio estrutural. Assim, seguindo o *Construal*, assumimos que o *parser*, ao encontrar a oração relativa, procede a uma associação e a apõe, rapidamente, com base na referencialidade do N1.

Duas questões principais motivaram a realização do experimento de rastreamento ocular: (1) a informação de gênero é acessada de imediato pelo *parser*? e (2) a referencialidade (do N1 e do N2) pode influenciar o processamento inicial da oração relativa? Os resultados permitiram esclarecer, em primeiro lugar, que não há um atraso no

acesso à informação de gênero, o que nos faz interpretar o efeito de gênero do experimento de leitura automonitorada como efeito *spillover*, e, em segundo lugar, que a referencialidade do N1 e do N2 influencia o modo de análise da oração relativa: quando apenas o N1 é referencial, a aposição não local seria favorecida, mas quando o N2 é referencial, a aposição local é que seria inicialmente favorecida.

Porém, os dois experimentos revelam, também, que o gênero do participio e a referencialidade do N2 podem eventualmente ser desconsiderados pelo *parser*, sobretudo no momento mais interpretativo do processamento sentencial, em favor de uma análise que parece ser a predominante no tipo de estrutura em estudo: a aposição ao N1. Ademais, a diferença entre os experimentos quanto aos resultados da referencialidade do N2 aponta para como o uso de diferentes metodologias experimentais pode ajudar a compreensão de um fenômeno. O modo de apresentação dos estímulos, neste sentido, poderia ter um papel importante: é legítimo cogitar a hipótese de que a não segmentação dos estímulos no rastreamento ocular pode favorecer mais a compreensão e a integração discursiva do N2 referencial do que a segmentação presente na leitura automonitorada, na qual isolamos o SN complexo, aumentando, possivelmente, a resistência ao N2 referencial e favorecendo uma decisão mais imediata de sua não introdução no modelo discursivo. Por fim, as diferentes medidas fornecidas pelo rastreamento ocular nos permitiram ver com mais clareza o que seria um efeito imediato ou tardio de nossas variáveis.

Em suma, fornecemos evidências de que o processamento de orações relativas associadas a SNs complexos do tipo “substância” é influenciado pela referencialidade dos núcleos que compõem tais SNs. Certamente, precisamos compreender melhor, por exemplo, a resistência ao N2 referencial vista, sobretudo, no primeiro experimento. Já comentamos que a relação sintático-semântica estabelecida entre o N1 e o N2 pode não favorecer a introdução do N2 referencial no modelo discursivo, e estamos trabalhando para alcançar uma explicação mais precisa.

Agradecimentos

Agradecemos ao professor Marcus Maia pelo seu apoio ao disponibilizar o equipamento e o ambiente físico necessários para a aplicação do experimento de rastreamento ocular. Agradecemos aos dois revisores anônimos pelas contribuições para a melhoria do texto.

Referências

- BRYLSBAERT, M.; MITCHELL, D. C. The failure to use gender information in parsing: a comment on van Berkum, Brown and Hagoort (1999). *Journal of Psycholinguistic Research*, Springer Link, v. 29, n. 5, p. 453-466, 2000. Disponível em: <<https://link.springer.com/article/10.1023/A:1005191308387>>.
- CHRISTIANSON, K.; HOLLINGWORTH, A.; HALLIWELL, J.; FERREIRA, F. Thematic roles assigned along the garden-path linger. *Cognitive Psychology*, Elsevier, v. 42, p. 368-407, 2001. <https://doi.org/10.1006/cogp.2001.0752>.
- CUETOS, F.; MITCHELL, D. C. Cross-linguistic differences in parsing: restrictions on the use of the Late Closure strategy in Spanish. *Cognition*, Elsevier, v. 30, p. 73-105, 1988. [https://doi.org/10.1016/0010-0277\(88\)90004-2](https://doi.org/10.1016/0010-0277(88)90004-2).
- DE VINCENZI, M.; DI DOMENICO, E. A distinction among phi-features: the role of gender and number in the retrieval of pronoun antecedents. *Rivista di Linguistica*, Laboratorio de Linguística Giovanni Nencioni, Pisa, v. 11, p. 41-74, 1999.
- FERREIRA, F. The misinterpretation of noncanonical sentences. *Cognitive Psychology*, Elsevier, v. 47, p. 164-203, 2003. [https://doi.org/10.1016/S0010-0285\(03\)00005-7](https://doi.org/10.1016/S0010-0285(03)00005-7).
- FERREIRA, F.; HENDERSON, J. M. Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, Elsevier, v. 25, p. 725-745, 1991. [https://doi.org/10.1016/0749-596X\(91\)90034-H](https://doi.org/10.1016/0749-596X(91)90034-H).
- FERREIRA, F.; PATSON, N. D. The ‘Good Enough’ Approach to Language Comprehension. *Language and Linguistics Compass*, Wiley Online Library, v. 1, p. 71-83, 2007.
- FERREIRA, F.; BAYLEY, K. G. D.; FERRARO, V. Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science*, Association for Psychological Science, v. 11, n. 1, p. 11-15, 2002.
- FODOR, J. D. Learning to parse? *Journal of Psycholinguistic Research*, Springer Link, v. 27, n. 2, p. 285-319, 1998. Disponível em: <<https://link.springer.com/article/10.1023/A:1023258301588>>.

FODOR, J. D. Psycholinguistics cannot escape prosody. In: SPEECH PROSODY CONFERENCE, 2002, Aix-en-Provence, France. *Proceedings of the Speech Prosody*, Aix-en-Provence: Laboratoire Parole et Langage, p. 83-88, 2002.

FRAZIER, L. *On comprehending sentences: syntactic parsing strategies*. 1979. Dissertation (Doctoral) - University of Connecticut, 1979.

FRAZIER, L. Sentence processing: a tutorial review. In: COLTHEART, M. (Ed.). *Attention and performance XII: the psychology of reading*. Mahwah, NJ: Erlbaum, p. 559-586, 1987.

FRAZIER, L. Exploring the architecture of the language-processing system. In: ALTMANN, G. (Ed.). *Cognitive models of speech processing: psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press, p. 409-433, 1990.

FRAZIER, L. Is 'Good Enough' parsing good enough? In: ARCURI, L.; BOSCOLO, P.; PERSOTTI, F. (Ed.). *Cognition and language: a long story*. Festschrift in honor of Ino Flores d'Arcais. Padua: Department of Psychology, University of Padua, p. 13-30, 2008.

FRAZIER, L.; CLIFTON, C. *Construal*. Cambridge: MIT Press, 1996.

FRAZIER, L.; CLIFTON, C. Construal: overview, motivation, and some new evidence. *Journal of Psycholinguistic Research*, Springer Link, v. 26, n. 3, p. 277-295, 1997. Disponível em: <<https://link.springer.com/article/10.1023/A:1025024524133>>.

FRAZIER, L.; RAYNER, K. Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, Elsevier, v. 14, p. 178-210, 1982. [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1).

GIBSON, E.; PEARLMUTTER, N.; CANSECO-GONZALEZ, E.; HICKOK, G. Recency preference in the human sentence processing mechanism. *Cognition*, Elsevier, v. 59, p. 23-59, 1996. [https://doi.org/10.1016/0010-0277\(95\)00687-7](https://doi.org/10.1016/0010-0277(95)00687-7).

GILBOY, E.; SOPENA, J.; CLIFTON, C. Jr.; FRAZIER, L. Argument structure and association preferences in Spanish and English complex NPs. *Cognition*, Elsevier, v. 54, p. 131-167, 1995. [https://doi.org/10.1016/0010-0277\(94\)00636-Y](https://doi.org/10.1016/0010-0277(94)00636-Y).

GRILLO, N.; COSTA, J. A novel argument for the universality of parsing principles. *Cognition*, Elsevier, v. 133, p. 156-187, 2014. <https://doi.org/10.1016/j.cognition.2014.05.019>.

GRODNER, D.; GIBSON, E.; WATSON, D. The influence of contextual contrast on syntactic processing: evidence for strong-interaction in sentence comprehension. *Cognition*, Elsevier, v. 95, p. 275-296, 2005. <https://doi.org/10.1016/j.cognition.2004.01.007>.

HAWKINS, J. A. *Definiteness and indefiniteness: a study in reference and grammaticality prediction*. London: Croom Helm, 1978.

HEIM, I. *The semantics of definite and indefinite noun phrases*. 1982. Dissertation (Doctoral) - University of Massachusetts, Amherst, MA, 1982.

HEMFORTH, B.; KONIECZNY, L.; SEELIG, H.; WALTER, M. Case matching and relative clause attachment. *Journal of Psycholinguistic Research*, Springer Link, v. 29, p. 81-88, 2000. Disponível em: <<https://link.springer.com/article/10.1023/A:1005176507878>>.

HEMFORTH, B.; FERNANDEZ, S.; CLIFTON, C.; FRAZIER, L.; KONIECZNY, L.; WALTER, M. Relative clause attachment in German, English, Spanish and French: effects of position and length. *Lingua*, Elsevier, v. 166, p. 43-64, 2015. <https://doi.org/10.1016/j.lingua.2015.08.010>.

JUST, M. A.; CARPENTER, P. A. A theory of reading: from eye fixations to comprehension. *Psychological Review*, Washington, v. 87, n. 4, p. 329-354, 1980. <https://doi.org/10.1037/0033-295X.87.4.329>.

KARIMI, H.; FERREIRA, F. Good-enough linguistic representations and online cognitive equilibrium in language processing. *The Quarterly Journal of Experimental Psychology*, Taylor & Francis, p. 1-28, 2015.

MAIA, M.; FINGER, I. Referencialidade e domínio temático na compreensão de orações relativas curtas e longas em português. *Linguística*, UFRJ, Rio de Janeiro, v. 3, n. 2, p. 249-278, 2007.

MAIA, M.; FERNÁNDEZ, E. M.; COSTA, A.; LOURENÇO-GOMES, M. C. Early and late preferences in relative clause attachment in Portuguese and Spanish. *Journal of Portuguese Linguistics*, John Benjamins, v. 6, p. 227-250, 2006. <https://doi.org/10.5334/jpl.151>.

RAYNER, K.; CARLSON, M.; FRAZIER, L. The interaction of syntax and semantics during sentence processing: eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, Elsevier, v. 22, p. 358-374, 1983. [https://doi.org/10.1016/S0022-5371\(83\)90236-0](https://doi.org/10.1016/S0022-5371(83)90236-0).

RIBEIRO, A. J. C. A abordagem good-enough e o processamento de frases do português do Brasil. *Veredas On-Line*, UFJF, Juiz de Fora, v. 2, p. 62-75, 2008.

RIBEIRO, A. J. C. Reanálise parcial de frases garden-path do PB. ENCONTRO NACIONAL DA ANPOLL, XXV., 2010. Belo Horizonte: UFMG, 2010. Trabalho apresentado.

RIBEIRO, A. J. C. Late closure e good-enough no processamento de frases garden-path do português do Brasil: evidências de eyetracking. *ReVEL*, v. 10, n. 18, p. 84-106, 2012.

SWETS, B.; DESMET, T.; CLIFTON, C. Jr.; FERREIRA, F. Underspecification of syntactic ambiguities: evidence from self-paced reading. *Memory & Cognition*, Springer Link, v. 36, n. 1, p. 201-216, 2008. Disponível em: <<https://link.springer.com/article/10.3758/MC.36.1.201>>.

WARREN, T; GIBSON, E. The influence of referential processing on sentence complexity. *Cognition*, Elsevier, v. 85, p. 79-112, 2002. [https://doi.org/10.1016/S0010-0277\(02\)00087-2](https://doi.org/10.1016/S0010-0277(02)00087-2).

**An ERP study of kind-denoting nouns
in the subject position in Brazilian Portuguese**

***Um estudo neurofisiológico de nomes que denotam tipo
na posição de sujeito em português brasileiro***

Aniela Improta França

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
aniela@gmail.com

Maurício Cagy

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
mauricio.cagy@gmail.com

Antonio Fernando Catelli Infantosi

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
mcagy@peb.ufrj.br

Resumo: A referencialidade é um tópico que sempre renova o interesse da linguística teórica e experimental. Isso porque as referências são recursos fundamentais das línguas naturais que estabelecem uma relação entre a língua no cérebro e as coisas do mundo lá fora. Este experimento investiga duas condições referentes a mecanismos de referencialidade: uma relaciona-se a um NP não específico na posição do sujeito e a outra se relaciona a um NP que denota tipo influenciado por um contexto especial de Conhecimento do Mundo. A comparação entre as duas condições nas suas versões congruente e incongruente revelou que a Condição do Conhecimento do Mundo apresentou um ERP (N_{400}) com latência mais curta. Já que as sentenças são percebidas em uma ordem linear e que o sujeito em português (língua SVO) vem na frente, aqui se propõe que o acesso ao sujeito pragmaticamente saliente na sua posição sentença-inicial abre, logo de início, uma moldura pragmática e pode

restringir a seleção no verbo-complemento, fazendo com que o merge verbo-objeto se dê de forma mais rápida. Isso pode explicar porque o N_{400} relacionado aos estímulos do Conhecimento do Mundo apareceu mais cedo do que na Condição de Controle. Adicionalmente, encontrou-se um ERP positivo em torno dos 600 ms apenas presente nos estímulos da Condição de Conhecimento do Mundo e ausente na Condição de Controle. A interpretação deste achado é que o pleno acesso ao conteúdo pragmático que está no sujeito só ocorre após o merge do complemento ao verbo.

Palavras-chave: hipótese sintaxe-primeiro; pragmática no processamento de sentenças; extração de potenciais relacionados a evento; N_{400} - P_{600} .

Abstract: Referring is a permanently hot topic in theoretical and experimental linguistics since references establish the fundamental relationship between parts of language in the brain and things out there in the world. This experiment investigates two conditions pertaining to reference mechanisms: one relates to a generic NP in the subject position and the other to an NP that is influenced by a special World Knowledge context. The comparison between the two conditions in their congruous and incongruous versions depicts a shorter latency for the N_{400} ERPs of the World Knowledge stimuli. Since linguistic input gradually reaches the reader from left to right, we propose that the access to a pragmatically salient subject in its sentence-initial position, by opening a window into World Knowledge, may restrict selections in the verb-complement merge more speedily than pragmatically unmarked subjects. This may explain why the N_{400} related to the World Knowledge stimuli appeared earlier than that of the Control Condition. Additionally, we found a later positive ERP around the 600ms only present in the World Knowledge stimuli and absent in the Control condition. The interpretation of this finding is that full access to the pragmatic contents of the subject only takes place after the verb complement merge.

Keywords: syntax-first hypothesis; pragmatics in sentence processing; event-related brain potentials; N_{400} - P_{600} .

Recebido em: 5 de janeiro de 2017

Aprovado em: 10 de abril de 2017

1 Introduction

Referring is a permanently hot topic in theoretical and experimental linguistics, since it takes the fundamental role of establishing a relationship between parts of language in the brain and things out there in the world.

Referring can take place in a number of ways. Typically, a referential expression (RE) is a Noun (N) that successfully picks out a definite referent in the world because the words in the expression and the way they are combined produce an accurate description of such referent, in such a way that the hearer of the expression can recognize the speaker's intention. (HAVILAND; CLARK, 2002; CLARK; SENGUL, 1979).

Notice, for instance (i) and (ii).

- (i) The dentist pulled the tooth easily
- (ii) Donald pulled the tooth easily

In (i), the noun *dentist* activates a specific piece of structured encyclopedic knowledge which is inextricably connected with that linguistic RE. Contrastingly, in (ii), out of a specific context or deitic situation, *Donald* is an RE that conveys poorer content, since it picks out any human being in the world to refer, just based on gender and number traces. Thus, this not-specific *Donald* exerts less attractive power in relation to *pull* and *tooth* than *the dentist*, because the latter is a kind-denoting RE.

But if the nouns in the subject position in (i) and (ii) are REs of different nature, it is important to verify how such contrasting nouns, respectively semantically rich and poor, might impact the online processing of these sentences.

At this point, holding minimal assumptions about on-line comprehension, it is safe to say that we parse an acoustic stream into discrete units that come into our ears incrementally. So, necessarily, part of the task is to identify the onset of each new syllable and combine syllables into words, phrases and sentence. As of the identification of the first words, there is profuse evidence that we take a very active, or better, proactive step in trying to anticipate incoming material, so that incoming items can, without a delay, be represented and incorporated into existing representations (SCHUBERTH; EIMAS, 1977; STANOVICH; WEST,

1981; ALTMANN; STEEDMAN, 1988; ALTMANN; KAMIDE, 1999; FEDERMEIER, 2007; ITO *et al.*, 2016; FREUNBERGER; ROEHM, 2016).

So, beyond the phonemic identification, during language processing, perceivers have a range of data at their disposal that they process on the fly: knowledge about the formal properties of language, semantic knowledge about contents and maybe even knowledge about pragmatics. But at what speed? What kicks in first? Does processing follow the linear acoustic or visual input order only or is it subject to a non-linear algorithm? Predictions depend on one's theoretical standing.

For instance, under a syntax-first account, inspired by the Minimalist framework (CHOMSKY, 1993), the assumption is that before any sentential meaning can be achieved, there is necessarily an intricate formal algorithm that secures structural hierarchy whirling around the vP, that includes the merge between the internal argument and the verb first. In this account, which is explicitly derivational, the sentence is built from the bottom up, bit by bit. Thus, since processing is incremental, in an SVO language, for a split of second, this syntax algorithm would make the listener ignore the noun that reached the ears and was perceived first, be it *the dentist* or *Donald*, to start work on the verb and its internal argument. Then the complete vP can license the external argument and merge with it. So, *pulled the tooth* is OK and needs an agent that can be *the dentist* or a human named *Donald*. Thus, the functional elements within the vP would be checked first, and the syntactic operations should be momentarily blind to the root content of the noun in the subject position regardless of its referential nature (CHOMSKY, 1998). In this view, the pragmatic frame deployed by the root of *dentist* would not be explored in the same time window as the one that applies to the vP, which is explored first (GOMES, 2014).

Contrastingly, taking a cognitivist, pragmatic approach to how sentential meaning is attained, Fillmore (1982, 2007) advances that REs like *dentist* are qualitatively special because they are semantically rich and they immediately activate a private history of that concept in the hearer's brain. This frame encompasses a coherent structure of related concepts, such as *pull* and *tooth*, that get combined by attraction of frequent use together. Thus, a frame would entail a net of facts about our world knowledge (WK) that can be retrieved from memory and that

refers to an entity kind, in this case that of a human being who works taking care of someone else's teeth.

Under the semantic-first account, the predictions would be that when processing starts, due to the pressure for fast linguistic decision-making, frugal heuristics takes place and the pragmatic frame deployed by the semantic content in the roots of *dentist*, *pull* and *tooth* is explored first. For these theories, the content of such roots is immediately inspected and suffers attraction, rejection and other surface effects that ultimately motivate syntax (KOS *et al.*, 2010; KIM; OSTERHOUT, 2005; KUPERBERG *et al.*, 2007; CHOW; PHILIPS, 2013).

The big question is to decide which account is closest to the neurophysiological reality of such cognitive computations. When are these pieces of knowledge triggered and what triggers them – syntax, semantics or pragmatics – to be used online while processing language in the brain? How do these pieces of knowledge interact toward the final understanding of the sentence? What constraint is satisfied first?

In order to investigate these complex issues affecting the rapid online integration of syntax, semantic and world knowledge information, we built an online investigation of Brazilian Portuguese sentences using a useful technique in neuroscience of language. It is the event-related brain potential (ERP): the extraction, from the electroencephalogram (EEG) recording of raw brain activity, of the clean electrical potentials specifically related with the brain activity one wants to study, in this case, language processing. The technique consists of time-locking a target linguistic unit to the corresponding EEG signal emitted by the participant's brain. With this time-locking procedure it became possible to pinpoint the exact instant that the target linguistic stimulus was presented to the volunteer and the exact time it provoked a reaction (the cortical response) related to that part of the stimulus (cf. KUTAS; HILLYARD, 1980; KUTAS; HILLYARD, 1984; KUTAS *et al.*, 1984, FRIEDERICI, 1995; LAGE, 2005; LI; YANG; HAGOORT, 2008).

For our basic comparison, we contrasted two conditions: Condition 1: World Knowledge (WK) sentences, like those in (i); and Condition 2: Control Sentences (CT) like in (ii). Thus, our objective is to time in milliseconds the impact that a kind-denoting noun in the subject position of our WK condition compared to a non-specific one of our CT condition will pose on the course of processing, just as it reaches the object.

We are embracing the syntax-first hypothesis that supports a single processing stream, syntax all the way down. (GOUVEA, 2010; GOMES, 2014; GOMES, FRANÇA, 2015; STOCKALL; MARANTZ, 2006; PHILLIPS; LEWIS, 2009; LAU *et al.*, 2006). If this hypothesis is correct, the special content of a kind denoting noun in the subject position should not hasten processing, since, at first, the semantic content of the noun in the subject position would not be inspected. As we shall see, this prediction bears interesting connections with neurophysiological markers for semantic congruence in the ERP literature and motivated the two additional conditions that we inserted in this test.

During the first thirty years of linguistic ERP research since the 80's, among the most commonly extracted ERP waves were those related to the semantic violations coinciding with the merge operation between the verb and its complement. The usual paradigms stimulate participants with sentences that contrast minimally as to congruence like in (iii) and (iv):

(iii) *John ate the sandwich yesterday*

(iv) *John ate the sandal yesterday*

Experimenters most commonly set the trigger for the stimulation point immediately after the verb and measure the ERP at that point, which relates to the exposure to the object. The literature shows that ERP peak amplitudes related to (iv) are higher than those related to (iii) and the difference between the two is statistically significant (KUTAS; HILLYARD, 1980; KUTAS, HILLYARD, 1984; HOLCOMB; NEVILLE, 1991, FRANÇA *et al.*, 2004). Usually the interpretation is that *sandals* is incongruous in the context of *eat* and therefore integration efforts result in augmented ERP amplitudes.

This very robust phenomenon, consistently tested across over 100 languages, can be described as an enlarged negative-going event-related potentials (ERP), peaking at around 400 ms post-stimulus: the N_{400} . The cortical reaction to such sentences has been widely described as a consequence of an increased difficulty in morpho-syntactic integration due to the fact that the semantic features of the noun cannot be integrated with the event the requirements set by the verb. The current study will also build on this legacy by proposing two additional conditions, presented in (iv) and (v) as Conditions 3 and 4. These are incongruous versions of Conditions 1 and 2.

- (v) The dentist pulled the thunder easily
- (vi) Donald pulled the thunder easily

Since ERPs are very sensitive to incongruous conditions, this well-tested congruence contrast for the N400 response was included in our test hoping it might interact with contextual or pragmatic data built-in via the subject, promoting an even more amplified ERP response that could discriminate better among conditions. Thus, crossing the two conditions relating to the subject semantics (world knowledge and control) with the two congruity versions (congruous and incongruous) we arrived at 4 conditions, that can be depicted in Table 1.

TABLE 1 – An example of a set of experimental stimuli in the four experimental conditions to be applied within-subject

Conditions	Sentences
Control Congruous (CTCC)	O Mauricio tirou o dente com facilidade <i>Maurício pulled the tooth easily</i>
Control Incongruous (CTIC)	O Mauricio tirou o trovão com facilidade <i>Maurício pulled the thunder easily</i>
World Knowledge Congruous (WKCC)	O dentista tirou o dente com facilidade <i>The dentist pulled the tooth easily</i>
World Knowledge Incongruous(WKIC)	O dentista tirou o trovão com facilidade <i>The dentist pulled the thunder easily</i>

We devised an experimental approach to compare sentences that are exactly the same in all their syntactic aspects, but that have semantically rich and poor REs in the subject position and congruous or incongruous nouns in the object position. Notice the stimuli in the four conditions have identical syntactic configurations: [subject [verb-object]], despite the fact that those in WK condition include world knowledge scenarios, and those in the Control condition that have a poor content RE in the subject position.

Thus, most crucially, exploring the predictions supposedly held by the Semantic-first account, if there is immediate inspection of the content of the noun in the subject position, CTCC should get a greater latency effect than WKCC. For instance, under this hypothesis, in stimulus 1, the world knowledge information inserted in the subject

dentist should narrow down the list of candidates for *pull* and the checking of the best candidate, *tooth*, should hasten the low amplitude N400 that will be formed after the verb.

Concerning the two incongruous conditions – CTIC and WKIC – the incongruence effect measured by the amplitude of the N400 should be enlarged in WKIC when compared with that in CTIC, due to the frustrated anticipation for a specific, focalized noun in the object position: in WKIC instead of the very predictable *tooth* the participant gets *thunder*. In CTIC, after *pull*, predictions include anything that can be pulled, so this unfocused, widespread activation of many candidates should soften the impact of incongruous *thunder* while enlarging its latency.

Contrastingly, under the perspective of Syntax-first approach, CTCC and WKCC should be exactly the same: the first noun comes in but it should not be fully inspected semantically. Then, when the verb *pull* comes in, there is anticipation for a large number of possible complements and among them, one is *tooth*. So, when the object comes and is semantically interpreted in both congruous conditions, it is judged as a compatible complement and the cortical reaction that derives from the integration of the verb and its complement should be a small amplitude N₄₀₀.

Again, under Syntax-first, CTIC and WKIC should be the same, since there is no special semantic contribution coming from the subject. So, the verb *pull* should elicit anticipation for a large number of possible complements and *thunder* is not among them. So, when the object comes in, it should be semantically interpreted and judged as incompatible with the verb and the integration efforts are supposed to cause a cortical reaction of a high amplitude N₄₀₀, the classical incongruence effect.

2 Materials and methods

This EEG-ERP test in its final configuration¹ was approved by the Ethics Committee of the University Hospital of the Federal University

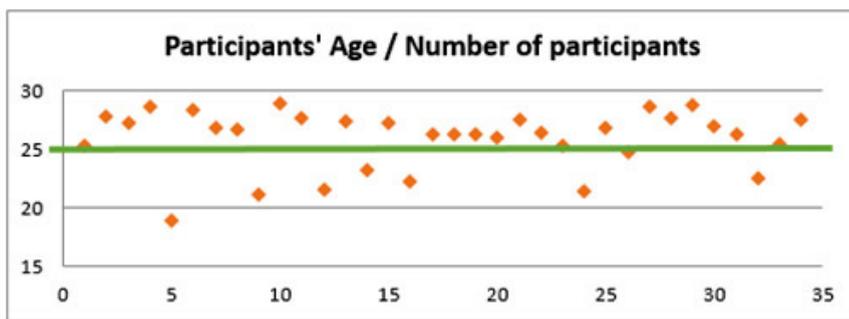
¹ An early version of this test was applied in 2006, and was registered as a work in progress in a non-commercial CD, internally distributed to the members of “GT de Psicolinguística da Anpoll”. The idea of the CD, organized by Marcus Maia and Ingrid Finger, was to increase the synergy among the group members of the GT by presenting their on-going, unfinished research. The version presented here includes 10 more participants, a revision of the stimuli and a full digestion of the findings.

of Rio de Janeiro, under the registration #113.05 (CEP – HUCFF/ UFRJ).² This committee also approved an informed consent form, that experimenters handed in to each participant prior to the beginning of the test. The form explained the methodology, its marginal risks and emphasized the fact that the participant may quit the test at any moment, and that no questions would be asked. The form was signed by all participants who took the test.

2.1 Participants and stimuli

While monitored by a 20 channel EEG, 34 right-handed participants (13 males) engaged in a computerized grammaticality judgment test. All participants were within the 18-39-age bracket (cf. Figure 1). Their mean age was 25.9 years (standard deviation, 3.0). All of them had a college education (complete or partial) and were previously screened for systemic diseases and for the current use of antidepressants and had normal or corrected-to-normal vision.

FIGURE 1 – Dispersion of participants' age (orange diamonds). Average (green line)



The experiment had a factorial design of 2x2, with 4 conditions. There were two independent variables related to the semantics of the noun in the subject position: non-specific, here named Control (CT), and kind-denoting, here named, World Knowledge (WK); each one had

² This project was submitted to the ethics committee by Professor Antonio Fernando Catelli Infantosi, a brilliant and most esteemed signal processing scholar from COPPE-UFRJ, co-author of this study, who died in 2016.

two levels: congruous (C) and incongruous (I). This design gave rise to four conditions: CTCC; CTIC; WKCC; WKIC. The stimuli in each of these conditions were designated to participants according to a within-subject distribution.

There were 160 experimental sentences divided among the four conditions. So there were 40 sets of stimuli like that in Table 1. Since there were four conditions, the within-subject distribution demanded we assigned participants to one of four test versions, so that participants did not see the same stimuli in more than one condition.

Each participant saw 40 experimental sentences, out of a total of 160 experimental sentences used in the test. Each participant was also assigned to 80 distractor sentences (40 congruous and 40 incongruous). All the 120 sentences each participant saw (40 experimental stimuli and 80 distractors), were displayed in a pseudo-random manner in order to perfectly control for sequential effects.

Both the subject and the object in all experimental sentences were introduced by a morphologically open DP (definite article) since proper nouns in Portuguese can be introduced by determiners. So this means that the control and the world knowledge conditions were formed by a DP_NP in the place of the subject. The occurrence of masculine and feminine determiners in the subject position was balanced among conditions.

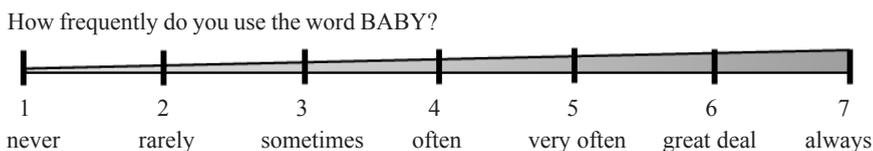
The distractors were pseudo-randomly mixed with the experimental stimuli and were formulated so as not to present only verb-complement selections as those in the target stimuli. For instance, while all experimental sentences included a transitive verb with agentive subjects, some distractors included intransitive verbs or light verbs and adverbials.³

³ Some of the distractor sentences were: (i) *O José trabalha muito no inverno* (he works a lot in the winter); (ii) *Ela sempre sorri as bananas verdes* (she always smiles the green bananas); (iii) *Os supermercados são caros* (the supermarkets are expensive); (iv) *Os policiais têm penas azuis* (the policemen have blue feathers). Sentences of type (i), with pronominal subjects and congruous semantics, were formulated to be distractors for congruous Control condition sentences. Sentences of type (ii), with pronominal subjects and incongruous semantics, were formulated to be distractors for incongruous Control condition sentences. Sentences of type (iii), with content subjects and congruous semantics, were formulated to be distractors for congruous WK condition sentences. Finally, Sentences of type (iv), with content subjects and incongruous semantics, were formulated to be distractors for incongruous WK condition sentences.

In selecting the stimuli, the criteria of word-length in syllables was taken into consideration. As to the frequency assessment, in the absence of a Brazilian Portuguese equivalent to a Cobuild Corpus and Frequency Levels, we resorted to a rating scale assessment, like the Likert scale of 7 levels, in which 1 corresponded to most infrequent and 7 to most frequent (DALMORO; VIEIRA, 2013).

We applied this test to 45 undergraduates of the Engineering school of the Federal University of Rio de Janeiro. We put together a collection of 100 nouns and we tested each one of them as to frequency before we made our stimuli. At the end, we selected only the nouns that achieved scores 6 or 7 by at least 75% of the participants. One item of the test follows as an illustration of the frequency test:

FIGURE 2 – Likert Scale presented to participants



Forty-three out of the 45 respondents (95%) rated the word *bebê* (*baby*) as 6 or 7 in the scale, since they reported using it “a great deal” or “always”. Thus, we selected BEBÊ as a suitable word to make one of the stimulus sentences of the test. Contrastingly, the word *quadro* (*painting*) was rated as 4 “used often” by 61% of the respondents, and was, thus, eliminated from the word collection used in the preparation of the experimental stimuli.

TABLE 2 – Additional examples of experimental stimuli in the four conditions

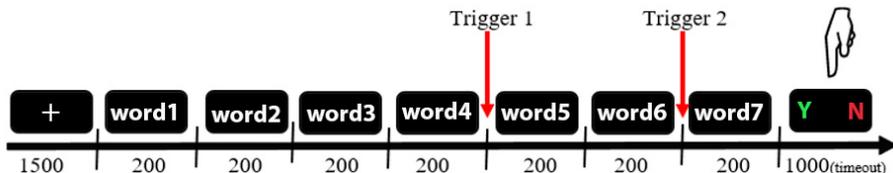
2	
CTCC	O Marcelo apagou o fogo rapidamente <i>Marcelo put out the fire rapidly</i>
CTCI	O Marcelo apagou a alma rapidamente <i>Marcelo put out the soul rapidly</i>
WKCC	O bombeiro apagou o fogo rapidamente <i>The fireman put out the fire rapidly</i>
WKIC	O bombeiro apagou a alma rapidamente <i>The fireman put out the soul rapidly</i>

3	
CTCC	A Angela curou o bebê totalmente <i>Angela healed the baby completely</i>
CTCI	A Angela curou o livro totalmente <i>Angela healed the book completely</i>
WKCC	A médica curou o bebê totalmente <i>The doctor healed the baby completely</i>
WKIC	A médica curou o livro totalmente <i>The doctor healed the book completely</i>
4	
CTCC	O Antonio puxou a arma na hora <i>Antonio pulled the gun instantly</i>
CTCI	O Antonio puxou a praia na hora <i>Antonio pulled the beach instantly</i>
WKCC	O soldado puxou a arma na hora <i>The soldier pulled the gun instantly</i>
WKIC	O soldado puxou a praia na hora <i>The soldier pulled the beach instantly</i>
5	
CTCC	A Fernanda tocou a valsa muito bem <i>Fernanda played the waltz very well</i>
CTCI	A Fernanda tocou o cheiro muito bem <i>Fernanda played the smell very well</i>
WKCC	O músico tocou a valsa muito bem <i>The musician played the waltz very well</i>
WKIC	O músico tocou o cheiro muito bem <i>The musician played the smell very well</i>

2.2 The stimulation protocol

Stimulus sentences were presented kinetically, word-by-word, on the computer screen, commanded by a script written in Presentation 0.5 (Neurobehavioral Systems, Albany, USA). Participants read test instructions on the screen followed by a warm-up drill that checked their comprehension of the protocol. After the warm-up, participants could receive additional coaching from the experimenter, in case doubts about the protocol still lingered. When ready, participants would start the congruity judgment test, that followed the timeline in Figure 3.

FIGURE 3 – The timeline of the experiment



Each word was centrally displayed on the monitor for 200 ms, formatted with white, 14-point, Times New Roman font over a black screen (170 monitor with a 1024 X 768 resolution). At the onset of the display of words 5 and 7, a trigger (square wave) was sent via parallel port to the computer that controls the experiment. This trigger is coupled to the waves sent by the EEG in a way to visually label them.

After the presentation of the last word of each sentence, subjects were to judge the stimulus for incongruence or congruence by pressing either the red N or green Y keys that are special keys that are placed on top of two regular keys in the computer keyboard, respectively to respond to the stimuli (*cf.* Figure 4).

FIGURE 4 – The keyboard with two specially identified keys, respectively for Y and N responses



This task allowed us to verify if participants were engaged in the experiments and were reading the sentences attentively. The high rate of correct responses in this off line test might indicate that participants were in fact distinguishing between congruous and incongruous stimuli. Response wait would time-out after 1000 ms. Following the event of judgment or time-out, a white fixation cross was displayed for 1500 ms before the first word of the next sentence was presented.

2.3 EEG Acquisition

EEG signal was continuously recorded during the whole experimental session from 20 monopolar derivations. Silver-tip electrodes were topographically positioned according to the International 10-20 System (Jasper, 1958), with averaged mastoid reference and ground at FPz.

Electrode impedance was controlled to normal values (for EEG, lower than 5 k Ω). Signal was amplified (gain = 18,000) and treated with low-pass (cut-off frequency of 32 Hz) and high-pass filtering (0.8 Hz). This filter allows the removal of all slow and constant electrical waves from the signals. All EEG derivations were digitized with a sampling frequency of 1000 Hz (12-bit analog-to-digital resolution) and were stored for off-line processing.

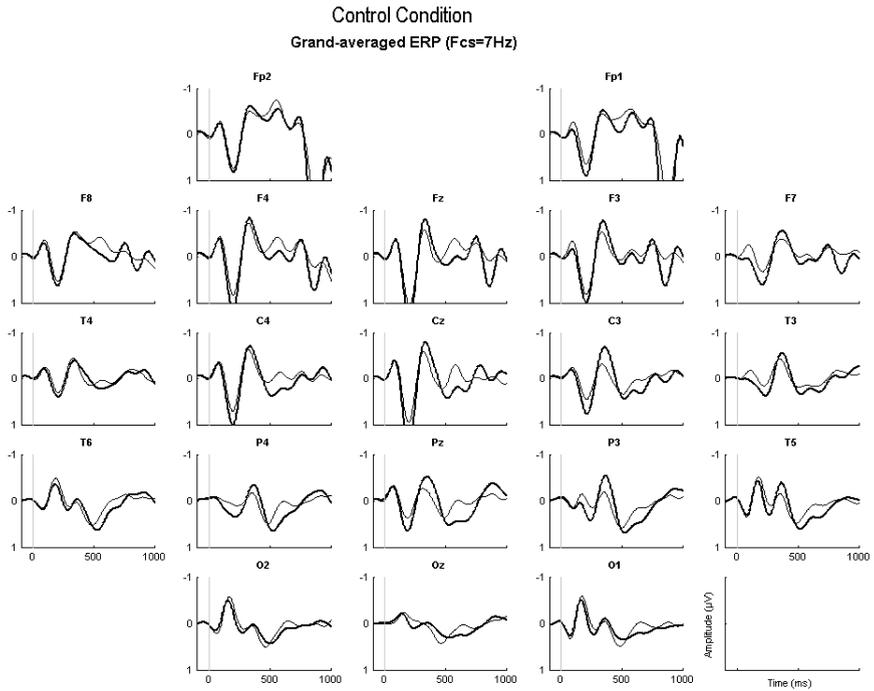
2.4 ERP Waveform Estimation

The multi-channel EEG digital processing of all experimental sentences was performed using Matlab version 5.2 environment (The MathWorks, Inc., MA). The original signal of each subject was segmented into epochs from 500 ms before up to 1500 ms after the onset of the target words. Then, an algorithm for artifact rejection was applied to each signal epoch. This algorithm consisted of comparing each EEG sample to a threshold, defined as 1.35 times the Root Mean Squared (RMS) value of an artifact-free individual EEG raw signal. The epochs that presented, within the time interval 0-800 ms, 10% of all samples or 5% of consecutive samples with absolute value above this threshold were discarded.

The ERP was then estimated by coherently averaging the epochs relative to congruous (or incongruous) EEG response for each topographic derivation of a subject. Hence, ERPs were time-locked to the onset of the stimulus-trigger for each of the two conditions (congruous and incongruous). A 100 ms pre-stimulus period was used to correct the baseline of each epoch. Individual ERPs were then low-pass filtered (cut-off frequency of 7 Hz, 2nd order Butterworth, applied bi-directionally for obtaining null phase frequency-response, i.e. no phase distortion, and consequently no change in latency). Figures 5 and 6 show the resulting grand-average of the 34 individual ERPs, from 100ms before up to 1000ms after the trigger word, where thick and thin lines refer to incongruous and congruous sentences respectively. Negative waves are plotted upward in accordance with ERP literature.

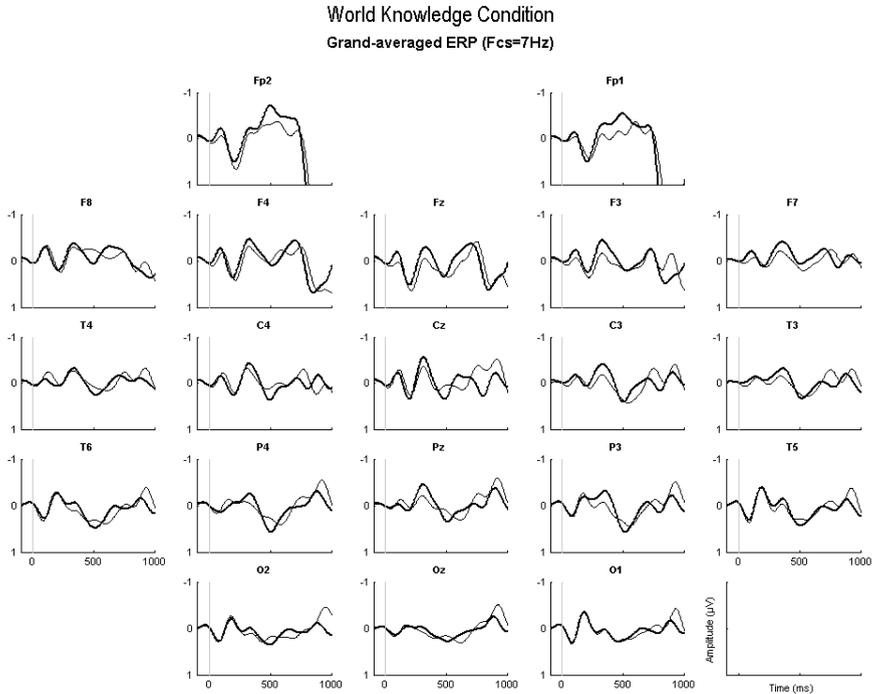
3 Results

FIGURE 5 – Grand-average ERPs from 34 right-handed subjects submitted to the Standard Condition



Note: EEG was extracted from 20 different cortical regions, in accordance with the International 10-20 System and with averaged mastoid reference (medial line: Fz, Cz, Pz and Oz; Fp2, F4, F8, C4, T4, P4, T6, O2 and their homologous). The thin line refers to ERPs resulting from congruous sentences and the thick line to those of the incongruous ones. Negativity is plotted upwards.

FIGURE 6— Grand-average ERPs from 34 right-handed subjects submitted to the Standard



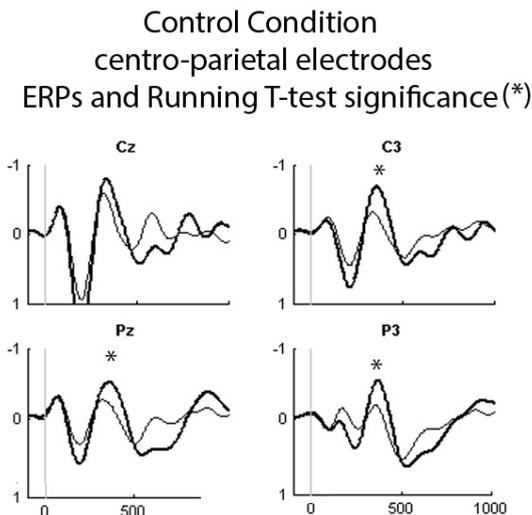
Condition. EEG was extracted from 20 different cortical regions, in accordance with the International 10-20 System and with averaged mastoid reference (medial line: Fz, Cz, Pz and Oz; Fp2, F4, F8, C4, T4, P4, T6, O2 and their homologous). The thin line refers to ERPs resulting from congruous sentences and the thick line to those of the incongruous ones. Negativity is plotted upwards.

3.1 Wave Morphology Reading

Since there are specific cerebral regions that respond to specific linguistic stimuli, the manipulation we introduced through the careful streamlining of the stimulus sentences in each condition affected mainly a specific cerebral region: the left centro-parietal derivations. To make interpretation of the ERPs easier to follow, we prepared Figures 7 and

8 that show the congruous – incongruous contrast in ERPs picked by electrodes that were positioned in this left centro-parietal region of interest.

FIGURE 7 – ERPs in the Control Condition in 4 derivations of interest

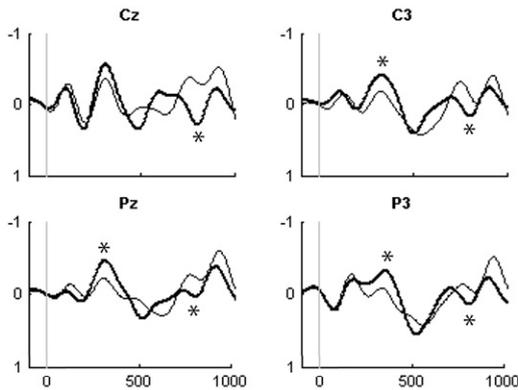


Legend: Thin line is congruous and thick line incongruous, the stars stand for statistically meaningful wave contrast.

In terms of latency, in the Control Condition (Figure 7), the amplitude of the incongruous N_{400} wave (thick line) is higher than that of the congruous one. Concerning the overall morphology, congruous and incongruous show similar shapes. Notice also that this condition yields basically one ERP within the 400-800 ms window.

FIGURE 8 – ERPs in the World Knowledge in 4 derivations of interest

World Knowledge Condition
 Centro-parietal Electrodes
 ERPs and Running T-Test significance(*)



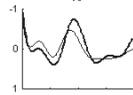
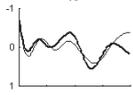
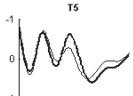
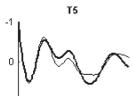
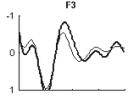
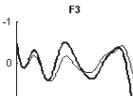
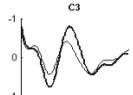
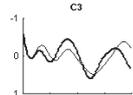
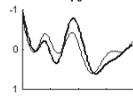
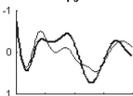
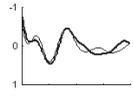
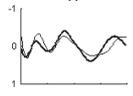
Legend: Thin line is congruous and thick line incongruous,
 the stars stand for statistically meaningful wave contrast.

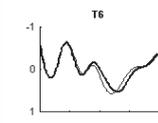
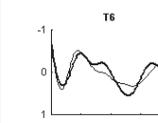
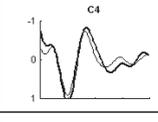
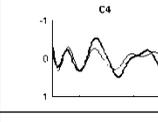
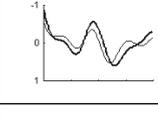
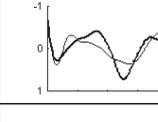
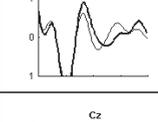
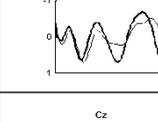
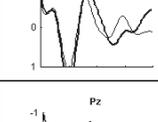
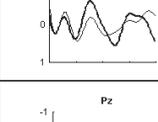
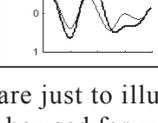
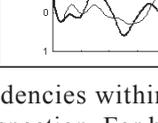
In terms of latency, in the World Knowledge Condition (Figure 8), the amplitude of the incongruous N_{400} wave (thick line) is higher than that of the congruous one. Concerning the overall morphology, congruous and incongruous show similar shapes. Notice also that this condition yields basically two ERPs within the 400-800 ms window. At around 400 ms, derivations with the largest amplitude make up an early N_{400} peak varying topographically from 300 to 375 ms. On the other hand, in the Control condition, the N_{400} exhibits a smaller topographical variability (range of 330-377 ms) (Figure 8). Beyond the N_{400} , in the WK condition we also find a positive potential (cf. Cz, Pz, P3), that deepens at around 560-620 ms, whose amplitude is considerably larger in the incongruous stimuli. This second wave is absent in the Control Condition. The amplitude of the incongruous N_{400} is also higher than that of the congruous one, but the ERP waveforms particularly in P3, Pz and P4 show noticeable differences.

3.2 Statistical Reading

Although there are considerable latency differences between the Control and WK N₄₀₀ ERPs (cf. Figures 7 and 8, Table 3), there is no evident difference in the N₄₀₀ latency between congruous and incongruous ERP within each condition. Therefore, no statistical comparison was performed related to N₄₀₀ latency. Nevertheless, a comparison of N₄₀₀ instantaneous amplitude between individual ERPs from incongruous (I) and congruous (C) sentences in each condition was carried out by means of the Running *t*-Test (HAGOORT *et al.*, 2004).

TABLE 3 – Intervals in which the waveforms for incongruous and congruous sentences differ significantly (between brackets).

Derivation	Control Latencies in ms	Control Waves	WK Latencies in ms	WK Waves
T3	[345-460]		[365-430] [715-790]	
T5	[505-725]		[370-410]	
F3	[315-430]		[255-420]	
C3	[315-440]		[310-435] [535-620]	
P3	[315-455]		[315-443] [765-800]	
T4	[410-440]		[625-670]	

T6	[405-465]		[315-355] [600-680]	
C4	[375-420]		[470-520] [755-800]	
P4	[360-445];		[315-380] [590-630]	
Fz	[335-420]		[270-310]	
Cz	[335-430]		[585-635]	
Pz	[315-450]		[300-365] [745-800]	

Note: The wave thumbnails are just to illustrate the ERP tendencies within the time interval, not supposed to be used for wave morphology inspection. For better visualization of the morphology of the waves, *cf.* Figures 4 and 5.

Using the differential waveform (*I-C*), Gaussianity could be assumed and the *t*-test was then applied sample-by-sample (each 5 ms) within the time interval of interest, that is from 200 to 800 ms. Two one-sided tests comparing (*I-C*) to zero, i.e. the null hypothesis of zero difference, were performed using a significance level of 2.5% for each side. At a given instant of time, if the null hypothesis is rejected for the positive side, one can infer that the resulting grand-average for incongruous sentences is statistically more positive or less negative than that for congruous sentences. The counterpart reasoning is applied to the instants of time for which the null hypothesis is rejected for the negative side.

Table 3 shows the intervals when the waveforms for incongruous and congruous sentences differ (between brackets). For example, in Pz, the time intervals [315-450] ms for Control condition and [300-365] ms for WK indicate more negativity for incongruous waveforms (thick lines) than for congruous (thin lines). Similar results, indicating that N400 is statistically more pronounced due to incongruence, are noted in most derivations except for Fp1, Fp2, F7, F8, T3 and C4 in the Control condition and O1 in the WK one. Further, an earlier N400 occurs in the WK if compared to Control condition. In Pz, the interval [745-800] ms for WK condition exhibits a more pronounced positivity for incongruous waveforms. The same kind of positivity is observed in other derivations except for Fp1, Fp2, Fz, F3, F7, T4, T5 and O2.

4 Discussion

The most interesting finding revealed by this experiment is that the sentences in the Control conditions yielded only one ERP, the N₄₀₀, while those in the World Knowledge conditions yielded two, respectively the N₄₀₀ and the P₆₀₀. In fact, other studies (LAGE, 2005; COWLES; KUTAS; KLUENDER, 2003; VAN BERKUM *et al.*, 2003; HAGOORT *et al.*, 2004) have found similar ERP results reporting an anterior negativity (N₄₀₀) coupled with P₆₀₀ effects in the establishment of anaphoric links involving world knowledge introduced by the subject. But we found no studies that streamlined conditions so as to properly compare a sentence with poor world knowledge content with the exact same sentence with an NP that deployed a rich pragmatic frame.

The syntax-first assumption that the course of the derivational process starts with the bottom merge does not completely conform with the results we found, since it cannot be applied to each and every sentence with a transitive verb. In our experiment, the World Knowledge stimuli that have a kind-denoting DP in the subject position, hastened the computational course. So it means that some information was indeed sensed linearly, as soon as the kind denoting subject entered the computation.

Nevertheless, the idea that the frame established by the subject would completely shape the general understanding of the sentence did not hold true either. Notice the contrast *pull the tooth Vs pulled the thunder*. This contrast was effective both for the Control and for the WK

conditions. In both conditions we have a large amplitude N_{400} relative to the incongruous version. But only in the WK condition is there a P_{600} relative to the incongruous condition. So, it is clear that the lexical, semantic and pragmatic components of the DP in the subject position of the WK condition are only fully digested at the end of the computation of the sentence, with the P_{600} . So, if information introduced by the subject, that comes in first, can only be sensed at 600 ms and information about the vP merge can be sensed at 400 ms, this reveals that there is precedence of some information of the vP merge in relation to some information stemming from the subject.

These facts are indicative that the linear online order of processing has to interact with the derivational phases of grammar that are available in the brain. So, in an itemized way, here are the conclusions.

a. Concerning the syntactic computation, in the WK condition, incongruence is established by the integration of the subject after the verb-object merge. That is, in this condition, our prediction is that incongruence is established later in the derivation, at the subject-predicate merge.

b. Concerning the Mental representation of contents, to judge the congruence of the control condition sentences, we have to know the semantic selectional properties of the verbs. For instance, *extinguish* selects an object with the semantic property of *fire*; *heal* selects an object with the semantic property of *living being*. In the incongruous version of the control condition (CTCI), semantic anomaly is established by and limited to the lower merge.

In the WK condition, incongruity judgments include information introduced by the subject merge, conforming to representations of world facts, including the notion, for instance, that *a dentist pulls tooth*, *a fireman extinguishes fire*. So, the incongruity in such cases stem from the subject merge and requires World Knowledge with special scenarios, which may encompass actual facts like treating a patient or putting out a fire or make-believe conventions like spinning a web for fictional characters like Spiderman, inserted in what has been termed a *Possible World Box* (NICHOLS; STICH, 2000).

A semantic anomaly like *eat shoes* produces an incongruous electric effect with a latency of about 400 ms and a large amplitude, the N_{400} , an effect that has been widely attested in the ERP literature and, in our work, was statistically indicated by means of a running *t*-test ($\alpha = 0.05$). This electric effect is said to result from difficulty of verb-

object integration: for instance *shoes* with its property [leather artifact] would not be selected by *eat*. The stimuli composed for experiments usually have the poorest content subjects possible, like a pronoun, any proper name or a non-specific kind (man, woman, boy etc), for instance. Thus, the subjects chosen for the Control Condition stimuli were clearly less restrictive than those picked for the WK condition. For instance, *dentist* is a fixed reference with defined objectives clothes, actions and origin while, *a man* or *Donald* can be any male human.

c. Concerning the contrast between the two conditions described above, the first finding we intend to discuss is the difference in latency between the first electrical wave in the Control condition and that in the WK one. Look at the N_{400} latencies in the Control condition and compare them to those in the WK condition, presented in Table 3. Note that the N_{400} related to the vP-object merge in the WK condition is earlier than that in the Control condition. How can we explain the fact that a verb-object merge involving special World Knowledge contents hastens the course of the derivation?

In the processing of Brazilian Portuguese, an SVO (subject-verb-object) word order language, prior to the verb-complement computation, there is the linear precedence of the subject. For the subject to be saved until syntactic computation integrates it into the structure being formed, it has to be stored in short term memory. For that, lexical access is required. And lexical access is automatic to allow whatever cognitive handling is needed, including storing.

If we look at the first NP as a transitorily independent expression, whose lexical access will guarantee memory storage of content for later use, we can suppose that the lexical content accessed is not completely innocuous to on-going syntactic computation and, in fact, affects the processing of the verb-complement constituent. If this is the case, the presence of a subject that recruits a non-default context hastens syntactic processing because it poses special requirements that restrict the VP internal selections.

d. Concerning the P_{600} , a second ERP that deepens at about 600 ms in incongruous sentences of the WK condition (cf. Table 3, WK derivations at T3, P3, T6, C4, P4, Fz, Cz and Pz). Our interpretation is that this positivity exclusive to this condition, is related to the repair efforts to integrate contents from kind-denoting DPs into the on-going subject-VP merge. We assume here that the integration of this kind of DP with the

vP constituent is a two-fold operation: first it prompts an early semantic access that accelerates the vP merge, but this early independent lexical access does not spare the later full computation required to merge the subject with the vP, which can be seen in the P₆₀₀ component.

We are hereby challenging the frame analysis that the subject-object semantic synergy spares the language user from making the regular syntactic computations to effect the vP-object merge operation. Our claim is that the array of NP candidates potentially licensed at the vP merge is much restricted in view of the access to a world-constructing non-default subject.

e. Concerning the higher wave amplitudes of the object merge coming from the congruous and incongruous sentences in the WK condition, a similar amplitude effect was reported by Hagoort *et alii.* (2004) in the gamma radiation related to stimuli requiring World Knowledge processing.⁴ Knowing that amplitude is a parameter related to a difficulty to perform morpho-syntactic integration (KUTAS; HILLYARD, 1984; FONTENEAU; FRAUENFELDER; RIZZI, 1998; HOLCOMB; NEVILLE, 1991), we propose that the pragmatic scope set by the subject DPs in the WK condition presented extra selectional restrictions to the verb, which ultimately overburdened the verb-object integration, causing the higher wave amplitudes captured at the derivations. Since the introduction of the subject shapes the causation mode in the VP (PYLKKÄNEN, 2002; MARANTZ, 2001), the retrieval of the subject DP requires encyclopedic contents, and so way more cognitive resources are demanded.

⁴ Hagoort *et alii* (2004) examined the wavelet band of theta and gamma radiation in relation to retrieval and usage of world knowledge and of knowledge of word meaning encoded in sets of stimuli with congruity violations. The authors found greater gamma electric activity as the oscillatory brain responses in relation to the world knowledge violations and greater theta activity in relation to the purely semantic conditions. Despite the fact that the former finding is compatible with our ERP findings, a direct comparison between the two experiments cannot be made since Hagoort *et alii.* used light verbs in the stimuli: *Dutch trains are yellow / white / sour*. Light verbs do not impose selectional restrictions confined to the vP level. As to latency, the authors explain that the wavelet transform cannot be a reliable measure of temporal resolution, thus, we cannot compare their latencies with ours.

5 Conclusion

This comparative ERP analysis of time-locked stimuli brought about an unexpected result: an N_{400} that is earlier in the WK Condition than in the Control Condition (running t -test, $\alpha = 0.05$). We also observed that, concerning congruity, the incongruous sentences in the WK condition generated more pronounced effects on the ERP morphology than those in the Control condition. These results can be properly interpreted if we take into account that access to the subject and its storing for future syntactic computation will entail a regular or modified array of verb semantic selection restrictions. Entering a non-trivial world knowledge space causes three effects in the waves locked to vP processing: N_{400} waves coming from the WK Condition are faster and more pronounced than those in the Control condition, the special semantic content of subjects in the WK condition impose a further integration, P_{600} , that can only take place after the verb is fully merged with its object (bottom-up order). The P_{600} might be related to the repair efforts to integrate special contents into the on-going subject-vP merge. This extended computation is not necessary for the Control condition because the special merge between verb and object is not added of any special world knowledge meaning.

Acknowledgments

Aniela Improta França gratefully acknowledges the financial support to her research (PQ 312079/2016-8) from CNPq, Brazil (National Counsel of Technological and Scientific Development).

References

- ALTMANN, G.; STEEDMAN, M. Interaction with context during human sentence processing. *Cognition*, Elsevier, v. 30, p.191-238, 1988.
- ALTMANN, G.; KAMIDE, Y. Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, v. 73, p. 247-264, 1999.
- CHOMSKY, N. A minimalist program for linguistic theory. In: ed. HALE, K.; KEYSER, S. *The view from building 20: Essays in honor of Sylvain Bromberger*. Cambridge, Mass: MIT Press, 1993. p. 1-52.

CHOMSKY, N. Minimalist inquiries: the framework. *MIT Working Papers in Linguistics*. Cambridge, MA: MIT 1998.

CHOW, W.-Y.; PHILLIPS, C. No semantic illusions in the “Semantic P600” phenomenon: ERP evidence from Mandarin Chinese. *Brain Research*, Elsevier, v. 1506, p. 76-93, 19 abr. 2013.

CLARK, H. H.; SENUGL, C. J. In search of referents for nouns and pronouns. *Memory and Cognition*, Springer, v. 7, p. 35-41, 1979.

COWLES, H. W.; KUTAS, M.; KLUENDER, R. Different ERP results from identical sentences: The importance of prior context in sentence processing. In: ANNUAL CUNY CONFERENCE ON HUMAN SENTENCE PROCESSING, 16th., Cambridge, 2003. Cambridge, MA.: Massachusetts Institute of Technology, 2003. Poster presented.

DALMORO, M.; VIEIRA, K. M. Dilemas na construção de escalas Tipo Likert: o número de itens e a disposição influenciam nos resultados. *Revista Gestão Organizacional*, UnoChapécó, v. 6, n. 3, p. 161-174, 2013.

FEDERMEIER, K. D. Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology*, Wiley Online Library, v. 44, p. 491-505, 2007.

FERREIRA, F.; CLIFTON, C. The independence of syntactic processing. *Journal of Memory and Language*, Elsevier, v. 25, p. 348-368, 1986.

FILLMORE, C. Frame semantics. In: The Linguistic Society of Korea (Ed.). *Linguistics in the Morning Calm: Selected Papers from SICOL-1981*. Seoul, Korea: Hashin Publishing Co., 1982. p. 111-137. [Reprinted in: GEERAERTS, D. (Ed.). *Cognitive Linguistics: Basic Reading*. Berlin: Mouton de Gruyter, 2006.]

FILLMORE, C. Valency issues in FrameNet. In: HERBST, T.; GÖTZ-VOTTELER, K. (Ed.). *Valency: Theoretical, Descriptive and Cognitive Issues*. Berlin: Mouton de Gruyter, 2007. p. 129-160.

FILLMORE, C.; BAKER, C. Frames approaches to semantics analysis. In: HEINE, B.; NARROG, H. (Ed.). *Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press, 2010. p. 313-340.

FONTENEAU, E.; FRAUENFELDER, U.; RIZZI, L. On the contribution of ERPs to study of language comprehension. *Bulletin Suisse de Linguistique Appliquée: Etudes en neurolinguistique*, Association Suisse de Linguistique Appliquée (ASLA), v. 68, p. 111-124, 1998.

FRANÇA, A.; LEMLE, M.; CAGY, M.; CONSTANT, P.; INFANTOSI, A. F. Discriminating among different types of verb-complement merge in Brazilian Portuguese: an ERP study of morphosyntactic sub-processes. *Journal of Neurolinguistics*, Elsevier, v. 17, p. 425-437, 2004.

FREUNBERGER, D.; ROEHM, D. Semantic prediction in language comprehension: evidence from brain potentials. *Language, Cognition and Neuroscience*, Taylor & Francis Online, v. 31, n. 9, p. 1193-1205, 2016. <http://dx.doi.org/10.1080/23273798.2016.1205202>.

FRIEDERICI, A. D. The time-course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, Elsevier, v. 50, p. 259-281, 1995.

GOMES, J. N. *Investigating the dissociation between N₄₀₀ and P₆₀₀ effects on the syntax semantics interface: an ERP study*. 2014. 183f. Tese (Doutorado) – Faculdade de Letras, da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2014.

GOMES, J.; FRANÇA, A. I. A técnica de ERP: Investigando a assimetria sujeito-objeto na interface sintaxe-semântica com EEG. *Letras de Hoje* (Online), PUCRS, v. 4, p. 152-163, 2015.

GOUVEA, A. C. *et al.* The linguistic processes underlying the P600. *Language and Cognitive Processes*, Taylor & Francis Online, v. 25, n. 2, p. 149-188, Feb. 2010. ISSN 0169-0965.

HAGOORT, P.; HALD, L.; BASTIAANSEN, M. PETERSSON, K. M. Integration of word meaning and world knowledge in language comprehension. *Science*, American Association for the Advancement of Science, v 304, p.438-441, April 2004.

HAVILAND, S. E.; CLARK, H. H. What's new? Acquiring new information as a process in comprehension. In: ALTMANN, G. T. M. (Ed.). *Psycholinguistics: Critical concepts in psychology*. London: Routledge, 2002. p. 512-521.

HOLCOMB, P. J.; NEVILLE, H. J. Natural speech processing: an analysis using event-related brain potentials. *Psychobiology*, Elsevier, v. 19, p. 286-300, 1991.

ITO, A.; PICKERING, M. J.; MARTIN, A. E.; NIEUWLAND, M. S. Predicting form and meaning: Evidence from brain potentials. *Journal of Memory and Language*, Elsevier, v. 86, p. 157-171, 2016. doi:10.1016/j.jml.2015.10.007.

KIM, A.; OSTERHOUT, L. The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, Elsevier, v. 52, n. 2, p. 205-225, Feb. 2005.

KOS, M.; VOSSE T.; VAN DEN BRINK, D.; HAGOORT, P. About Edible Restaurants: Conflicts between Syntax and Semantics as Revealed by ERPs. *Frontiers in Psychology*, Fontieres Media, v. 1, p. 1-11, Jan. 2010.

KUPERBERG, G. R.; KREHER, D; SITNIKOVA, T.; CAPLAN, D.; HOLCOMB, P. The role of animacy and thematic relationships in processing active English sentences: evidence from event-related potentials. *Brain and Language*, Elsevier, v. 100, n. 3, p. 223-237, mar. 2007.

KUTAS, M.; HILLYARD, S. A. Reading senseless: brain potentials reflect semantic incongruity. *Science*, American Association for the Advancement of Science, v. 207, p. 203-205, 1980.

KUTAS, M.; HILLYARD, S. A. Brain potentials during reading reflect word expectancy and semantic association. *Nature*, v. 307, Springer Nature, n. 5947, p. 161-163, 1984.

KUTAS, M.; LINDAMOOD, T. E.; HILLYARD, S. A. Word expectancy and event-related brain potentials during sentence processing. In: KORNBLUM, S.; REQUIN, J. (Ed.). *Preparatory studies and processes*. Hillsdale, NJ: Erlbaum. 1984. p. 217-237.

LAGE, A. C. Aspectos neurofisiológicos do processamento de sentença em português. 2005. Tese (Doutorado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

LAU, E. Lau; STROUD, C.; PLESCH, S.; PHILLIPS, Colin. The role of structural prediction in rapid syntactic analysis. *Brain and Language*, Elsevier, v. 98, n. 1, p. 74-88, Jul. 2006.

LI, X.; YANG, Y.; HAGOORT, P. Pitch accent and lexical tone processing in Chinese discourse comprehension: an ERP study. *Brain Research*, Elsevier, v. 1222, 192-200, 2008. doi:10.1016/j.brainres.2008.05.031

MARANTZ, A. *Words*. (MS Word document). Handout from WCCFL XX, 2001.

NICHOLS, S.; STICH, S. A cognitive theory of pretense. *Cognition*, New York, Elsevier, v. 74, n. 2, p. 115-147, 2000.

PHILLIPS, C.; LEWIS, S. Derivational Order in Syntax: Evidence and architectural consequences. In: CHESI, C. (Ed.). *Directions in Derivations*. Amsterdam: Elsevier, 2009. v. 2, p. 427-459.

PYLKKÄNEN, L. Introducing Arguments. 2002. Thesis (Ph.D.) - MIT, Cambridge, MA., 2002.

SCHUBERTH, R. E.; EIMAS, P. Effects of context on the classification of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, American Psychological Association, v. 3, p. 27-36, 1977.

STANOVICH, K. E; WEST, R. F. The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance*, American Psychological Association, v. 7, p. 658-672, 1981.

STOCKALL, L.; MARANTZ, A. A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, John Benjamins, v. 1, n. 1, p. 85-123, Jan. 2006.

VAN BERKUM, J.; BROWN, C. M.; HAGOORT, P.; ZWITSERLOOD, P. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, Wiley Online Library, v. 40, p. 235-248, 2003.

VAN PETTEN, C.; KUTAS, M. Influences of semantic and syntactic context in open- and closed-class words. *Memory & Cognition*, Springer, v. 19, p. 95-112, 1991.

Referência genérica em SNs singulares: uma abordagem cognitivista experimental

Generic reference in singular NPs: a cognitive grammar experimental approach

Diogo Pinheiro

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
diogopinheiro@letras.ufrj.br

Lilian Ferrari

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
lilianferrari@uol.com.br

Maria Clara Pimenta

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
maclarapimenta@yahoo.com.br

Joabe de Souza

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
joabesouza@hotmail.com.br

Márcia Viegas

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
marcia.viegas@uol.com.br

Wanderson Lucas Ohenes

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
wlohenes@gmail.com

Resumo: Este artigo se debruça sobre dois padrões de SNs singulares genéricos do português brasileiro – com artigo definido e com determinante zero – com um objetivo duplo: de um lado, caracterizar

as estruturas conceptuais associadas a cada um desses padrões; de outro, testar experimentalmente a validade dessa caracterização. À luz da Gramática Cognitiva (LANGACKER, 1987, 1991, 2009, 2013), assumimos que ambos os padrões designam um tipo (em oposição a uma instância particular) inserido em um domínio conceitual abstrato (em oposição ao espaço físico). Com base nisso, sugerimos que (i) o esquema com artigo pressupõe a disponibilidade prévia do domínio de instanciação e evoca a conceitualização implícita de tipos não perfilados, ao passo que (ii) o esquema com determinante zero não pressupõe acesso prévio ao domínio de instanciação e não faculta a conceitualização de tipos não perfilados. Para testar essas hipóteses, foi realizado um experimento de julgamento de aceitabilidade no qual trinta estudantes de graduação avaliaram sentenças contendo SNs singulares genéricos definidos em duas condições: com sequência contrastiva (condição plenamente compatível com a hipótese (i)) e sem sequência contrastiva (condição menos compatível com a hipótese (i)). Os resultados revelaram a existência de diferença significativa na distribuição dos graus de aceitabilidade entre as duas condições ($p = .043$), fornecendo evidências em favor da proposta desenvolvida aqui.

Palavras-chave: Gramática Cognitiva; abordagem experimental; sintagma nominal genérico; português brasileiro.

Abstract: This paper investigates generic singular NPs in Brazilian Portuguese, comparing patterns containing a definite article with those that display a zero determiner. The goal is twofold: to characterize the conceptual structures associated with these predications and to experimentally validate this characterization. Within the framework of Cognitive Grammar (LANGACKER, 1987, 1991, 2009, 2013), we assume that both syntactic patterns designate a type (as opposed to a particular instance) which inhabits an abstract conceptual domain (as opposed to physical space). Given this assumption, we argue that: (i) the definite schema implies the previous availability of the domain of instantiation and implicitly conjures up unprofiled types, whereas (ii) the zero determiner schema does not presuppose previous access to the domain of instantiation and does not evoke unprofiled types. To test these hypotheses, an acceptability judgement test was carried out where 30 undergraduate students rated sentences containing definite generic NPs in two conditions: with a contrastive sequence (fully compatible

with hypothesis (i)) and without a contrastive sequence (less compatible with hypothesis (i)). The results revealed the existence of a significant difference in the distribution of acceptability ratings in each condition ($p = .043$), thus providing evidence for the conceptual approach advanced here.

Keywords: Cognitive Grammar; experimental approach; generic noun phrase; Brazilian Portuguese.

Recebido em: 10 de dezembro de 2016

Aprovado em: 26 de abril de 2017

*Quem vai pro exterior da favela sente saudade
O gringo vem aqui e não conhece a realidade
Vai pra Zona Sul pra conhecer água de coco
E o pobre na favela passando sufoco
(Cidinho e Doca – Rap da Felicidade)*

1 Introdução

A semântica dos sintagmas nominais genéricos tem merecido atenção especial na literatura, principalmente a partir de estudos em semântica formal (KRIFKA *et al.*, 1995). Em especial, todas as línguas parecem disponibilizar sintagmas singulares genéricos, mas divergem quanto às restrições semânticas e sintáticas impostas a seus usos. Assim, enquanto línguas não-relacionadas, como as germânicas e românicas, admitem o uso de SNs definidos com interpretação genérica (“o gato é voluntarioso”), o português brasileiro (PB) possui a peculiaridade, mesmo entre as línguas românicas, de contar com o singular genérico com determinante zero (“gato é voluntarioso”), conforme descrito por Müller (2002a, 2002b).¹

¹ A expressão “singular genérico com determinante zero” é empregada por Langacker (1991), no âmbito da sua *Cognitive Grammar*, com propósito meramente descritivo. É nesse sentido também que o empregamos aqui, já que a *Cognitive Grammar* é o quadro teórico no qual este trabalho se insere. Neste artigo, portanto, o termo “determinante zero” não pressupõe a hipótese de que exista, “na sintaxe”, a projeção de um determinante não realizado fonologicamente. Vale lembrar que esse tipo de

Dentre as vertentes teóricas que compõem a Linguística Cognitiva, a Gramática Cognitiva (GC) é aquela que apresenta a proposta mais detalhada para a caracterização semântica dos sintagmas nominais (LANGACKER, 1987, 1991). Essa proposta envolve a caracterização do significado dos SNs em termos de conceptualização. Tomando a abordagem langackeriana como ponto de partida, o presente trabalho busca investigar a diferença entre dois tipos de SNs singulares genéricos do PB – o esquema com artigo definido (“o gato é voluntarioso”) e o esquema com determinante zero (“gato é voluntarioso”) – com o objetivo de associar a organização interna dessas predicções a estruturas conceptuais específicas.

Na esteira da GC, defenderemos aqui que ambos os esquemas são representados em um Espaço de Tipo (ET), caracterizado como um domínio conceptual abstrato. Com base nisso, proporemos que os SNs genéricos estabelecem diferentes articulações entre o ET e outro domínio conceptual relevante, o Espaço Discursivo Corrente (EDC), definido como o espaço mental que inclui os elementos e as relações compartilhados por falante e ouvinte em cada momento do fluxo discursivo. Especificamente, sustentamos que o emprego de um SN singular genérico com artigo definido pressupõe a disponibilidade prévia do ET no EDC, ao passo que o mesmo não é verdade para o padrão com determinante zero.

Como se observa, esse tipo de abordagem envolve sugestões bastante específicas quanto ao tipo de *representação mental* associado ao emprego de SNs singulares genéricos. É preciso reconhecer, no entanto, que tratamentos dessa natureza abrem espaço para um questionamento importante: como é possível assegurar a *realidade psicológica* das descrições propostas? Com efeito, a insistência inicial da Linguística Cognitiva em análises estritamente linguísticas fundadas na intuição do analista levou alguns pesquisadores a questionar se o campo “faz jus ao seu nome” (PEETERS, 2001). Essa crítica, é preciso salientar, não poupou a GC langackeriana: algumas vozes, ainda que relativamente isoladas,

interpretação é inteiramente estranho à arquitetura gramatical postulada em modelos construcionistas (como é o caso da *Cognitive Grammar*), nos quais as representações fonológica e semântica não são “lidas” ou “interpretadas” a partir de uma estrutura sintática prévia (sobre esse ponto, ver Michaelis (2013)).

têm questionado a realidade psicológica de parte de suas propostas (BROCCIAS; HOLMANN, 2007; HOLLMANN, 2013).

Em face dessas preocupações, e em consonância com o que tem sido chamado de “guinada empírica” em Linguística Cognitiva (STEFANOWITSCH, 2011), propomo-nos neste trabalho a testar experimentalmente a validade das nossas hipóteses teóricas. Para isso, realizamos um experimento de julgamento de aceitabilidade no qual falantes nativos do PB foram solicitados a avaliar, em uma escala Likert de cinco pontos, a naturalidade de um conjunto de sentenças contendo SNs singulares genéricos. Ao recorrer a esse tipo de expediente para verificar a plausibilidade de certas postulações teóricas, esperamos poder contribuir para disseminar, também na comunidade de linguistas cognitivistas brasileiros, o interesse pelo recurso a métodos empíricos, e em particular experimentais.

O trabalho está organizado em três seções principais. Na seção 2, são apresentados os pressupostos teóricos que fundamentam a pesquisa, detalhando-se premissas básicas da GC bem como sua proposta de caracterização dos nomes e sintagmas nominais. Na sequência, a seção 3 apresenta uma proposta de descrição cognitivista – sob a ótica da GC – do aporte conceptual associado aos dois esquemas de SNs singulares genéricos investigados aqui. Por fim, a seção 4 descreve o experimento realizado e discute seus resultados.

2 Gramática Cognitiva, sintagma nominal e *grounding*

Nesta seção, enfocam-se os pressupostos teóricos que nortearam a pesquisa. Após uma breve apresentação da GC (LANGACKER, 1987, 1991, 2009, 2013), detalhamos a proposta de Langacker para a caracterização dos nomes e sintagmas nominais, com especial atenção aos SNs singulares genéricos. Por fim, enfocamos o processo de *grounding* associado a sintagmas nominais definidos e indefinidos.

2.1 Breve caracterização da Gramática Cognitiva

A Gramática Cognitiva, quadro teórico desenvolvido por Ronald Langacker (1987, 1991, 2009, 2013), tem como proposta fundamental a premissa de que a gramática, ao invés de ser autônoma, reside em padrões esquemáticos de estrutura conceptual e simbolização. Sob essa ótica, o

termo *predicação* é usado para designar o significado de uma expressão, independentemente de sua abrangência. Tal significado, é importante ressaltar, não diz respeito apenas ao conteúdo conceptual, mas também ao modo particular como esse conteúdo é construído. A noção de *construal* busca apreender esse fenômeno, referindo-se à habilidade do falante de conceber e retratar uma mesma situação de formas alternativas.²

Tendo em vista que o significado é concebido como interação entre conteúdo conceptual e *construal*, a GC adota o termo *domínio* para se referir, de modo uniforme, ao conteúdo. Sendo assim, considera-se que uma expressão evoca um conjunto de domínios cognitivos, coletivamente denominados *matriz*, como base para o significado. Por exemplo, a expressão “copo d’água” pode evocar os seguintes domínios de experiência: espaço, o conceito de água, o conceito mais esquemático de líquido (imane em água), a concepção de um contêiner com líquido, noções de volume, etc. Por um lado, a lista de domínios evocados por uma expressão não é exaustiva, e depende sempre do propósito comunicativo do falante; por outro, pode-se estabelecer uma distinção entre *domínio básico*, que é não derivável ou analisável a partir de outras concepções (ex. Espaço, tempo, espectro cromático, espectro sonoro, temperatura, gosto, cheiro, etc.) e *domínio não-básico*, que pode ser reduzido a outras noções (ex. Instâncias de experiência imediata de tipo sensorial, emotivo ou motor/sinestésico, como a sensação de medo; produtos abstratos de operações intelectuais, como os conceitos de JUSTIÇA, VERTEBRADO, etc.). Os domínios não-básicos tendem a se organizar em hierarquias, de modo que uma concepção relativa a um determinado nível pressupõe e incorpora uma ou mais concepções de nível mais básico.

Embora reconheça que nem todos os significados são baseados em percepção visual, a GC recorre à metáfora visual para classificar as várias facetas do *construal*. Essas facetas incluem as dimensões de *especificidade*, *focalização* e *proeminência*, as quais passamos a descrever.

A *especificidade* diz respeito ao nível de precisão com que uma situação é retratada. Assim, um determinado animal pode ser descrito como “roedor”, mas também como “esquilo”, o que implica maior grau

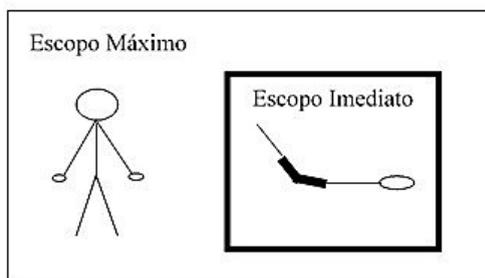
² Optamos por não traduzir o termo “construal” por se tratar de termo técnico cujas possíveis traduções (interpretação, construção) levariam a termos já utilizados na literatura linguística com acepções diferentes.

de especificidade. Nesse caso, o termo “roedor” é mais *esquemático* que “esquilo” – ou seja, considera-se que a caracterização “roedor” pode ser instanciada por caracterizações mais específicas, cada uma delas servindo para *elaborar* especificações mais detalhadas (roedor → esquilo → esquilo marrom → esquilo marrom grande).

A *focalização* é um aspecto do *construal* que inclui a seleção de conteúdo conceptual para apresentação linguística, como é o caso da organização em termos de *figura e fundo* (*foreground vs. Background*). Em termos mais gerais, todas as expressões evocam conhecimento pressuposto (*background knowledge*) como base para a interpretação. Por exemplo, para interpretar a sentença “Deixei meu carro no estacionamento rotativo”, é preciso acessar o conhecimento cultural referente a esse tipo de estacionamento, sem o qual o ouvinte poderia pensar que o carro foi abandonado no referido local. Além da organização figura vs. Fundo, a focalização inclui a extensão que uma expressão recobre no domínio acessado. Para cada domínio de uma matriz, uma expressão tem um *escopo* que consiste na sua cobertura daquele domínio, a qual será sempre limitada (*bounded*) em sua extensão. Assim, o termo *copo* evoca uma certa extensão espacial para a especificação de sua forma característica (mas não todo o universo), o verbo *tropeçar* requer o acesso mental a um período de tempo suficientemente longo para que o evento ocorra (e não a eternidade), e assim por diante.

Vale notar que o escopo, sendo uma questão de seleção, pode ser também organizado em termos de figura/fundo. Em determinados casos, é preciso distinguir entre *escopo máximo* de uma expressão em determinado domínio (ou seja, toda a extensão recoberta pela expressão) e o *escopo imediato*, mais limitado e diretamente relevante para um propósito particular. Usando-se uma metáfora teatral, o escopo imediato também pode ser denominado *onstage region* (região relativa ao palco, para a qual se dirige a atenção visual). Para ilustrar esse fenômeno, Langacker (2013, p. 63) discute a palavra “cotovelo” (*elbow*), que tem como um domínio bastante central de sua matriz a concepção de corpo humano. O autor nota, contudo, que o corpo não é um todo indiferenciado, mas uma estrutura composta de diferentes partes. Assim, o cotovelo é concebido, em primeiro lugar, como parte do braço. Há, portanto, uma hierarquia conceptual (CORPO HUMANO > BRAÇO > COTOVELO) em que BRAÇO funciona como escopo imediato e CORPO HUMANO, como escopo máximo:

FIGURA 1 – Escopo máximo e imediato para COTOVELO



Como mostra a Figura 1, a partir do escopo máximo, a expressão “cotovelo” seleciona um determinado conteúdo conceptual para colocar em *proeminência* (isto é, *onstage* no escopo imediato). Esse conteúdo selecionado constitui uma estrutura particular que será denominada *perfil* (*profile*). Por isso, diz-se que, no escopo imediato representado acima, a estrutura COTOVELO está *perfilada* (*profiled*).

2.2 Caracterização de nomes e sintagmas nominais

Valendo-se do instrumental teórico desenvolvido para tratar do significado das expressões linguísticas, Langacker (1991) enfoca as estruturas nominais, defendendo que a função semântica (e não o caráter estrutural) é o fator crítico para entender sua organização interna. Com base na proposta de que um nome perfila uma *coisa* (*thing*), definida como uma região (conjunto de entidades interconectadas) em um determinado domínio, o autor parte para a distinção das predicções nominais em termos de suas propriedades conceptuais específicas. Em especial, investiga a distinção entre nomes simples (“gato”, “telhado”, “vinho”) e SNs (“aqueles três gatos”, “um telhado de madeira”, “o vinho branco”), propondo que os primeiros exercem a função de *tipo*, enquanto os segundos caracterizam *instâncias* de um tipo.

2.2.1 Instanciação: tipo vs. Instância

Para Langacker (1991, p.55), a distinção entre tipo e instância apresenta pontos em comum com a distinção entre *extensão* e *intenção*. A extensão de um termo é o conjunto de objetos que podem ser designados

por ele em um dado mundo: a extensão de “gato”, por exemplo, é o conjunto de gatos. A intenção de um termo, por outro lado, é descrita como uma função das palavras para suas extensões, ou seja, um tipo de caracterização que identifica precisamente os objetos que constituem a extensão do termo no mundo.

Apesar dos pontos em comum, entretanto, o autor aponta diferenças importantes entre essas dicotomias. Em especial, os constructos *tipo* e *instância* correspondem a estruturas conceptuais e enfocam o modo pelo qual situações concebidas são retratadas linguisticamente. Sob a ótica da GC, “tipo” e “instâncias” equivalem a *conceito* e *entidades conceptuais*, respectivamente. Assim, a instância designada por “o gato” não equivale nem à referência da expressão nominal (tomada normalmente como um objeto no mundo), nem à extensão de “gato” (o conjunto de todos esses objetos). Isso porque o polo semântico de um SN é tratado como uma conceptualização, que pode designar até mesmo uma instância não-específica ou não-referencial (ex. “Ele queria encontrar *um apartamento* próximo ao trabalho, mas não havia *nenhum apartamento* à venda na região”).

Para elucidar o contraste entre *concepção de tipo* e *concepção de instância*, Langacker (1991) propõe a noção de *domínio de instanciação*, definido como o domínio no qual se considera que uma entidade reside ou tem sua manifestação primária. Por exemplo, TEMPO é o domínio de instanciação para eventos, enquanto ESPAÇO é o domínio de instanciação para substância material. Mais precisamente, o domínio de instanciação é o domínio no qual a localização de uma entidade é suficiente para estabelecê-la como uma instância de uma categoria distinta de outras instâncias potenciais.

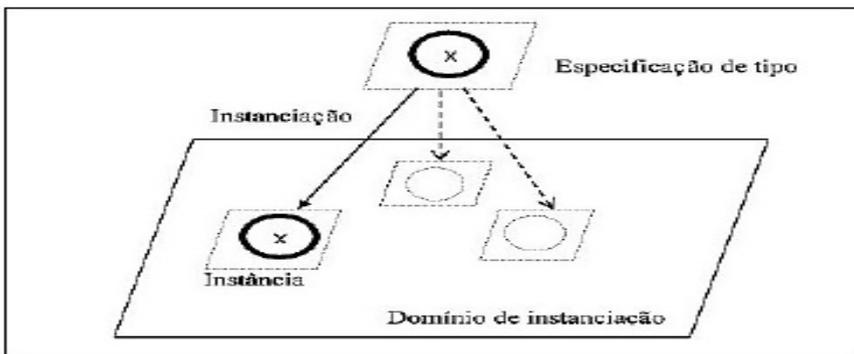
As concepções de tipo e instância compartilham a propriedade de perfilar uma entidade no domínio de instanciação. Por exemplo, tanto “gato” quanto “o gato” evocam e perfilam a concepção de uma criatura peluda que ocupa o espaço físico e tem uma determinada forma nesse domínio. Sendo assim, Langacker (1991) propõe um fator adicional para dar conta da distinção: *instâncias (mas não tipos) apresentam uma localização particular no domínio de instanciação*. Isso significa que uma especificação de tipo usa o domínio de instanciação apenas para descrever uma propriedade da entidade designada (como sua forma ou extensão temporal), ao passo que uma especificação de instância concebe o domínio como tendo extensão suficiente para sustentar manifestações

simultâneas de entidades múltiplas, caracterizando a entidade perfilada como situada no âmbito dessa extensão em uma localização específica em contraste com outras localizações possíveis. Nas palavras do autor:

Pode ser útil imaginar uma especificação de tipo como flutuando pelo domínio de instanciação, sem estar presa a nada, com o potencial de se manifestar em qualquer lugar dentro desse domínio. Esse potencial é concretizado, e uma concepção de instância é obtida, quando a especificação é ancorada em um ponto particular (LANGACKER, 1991, p. 57).³

Nos termos da definição proposta, a especificação de instância é assim representada:⁴

FIGURA 2 – Especificação de instância



³ “It may be helpful to imagine a type specification as floating about unattached through the domain of instantiation, with the potential to be manifested anywhere within it. This potential is realized, and an instance conception obtained, when the specification is anchored at a particular spot.”

⁴ Deve-se destacar que a noção de instanciação difere das noções de *grounding* e referência. O *grounding*, a ser detalhado mais adiante, pressupõe que uma instância foi estabelecida e acrescida de uma indicação relativa aos participantes do evento de fala; a instanciação em si não fornece essa indicação. Além disso, o efeito de uma predicação ancorada no *ground* (*grounding predication*) pode ser o de negar o status referencial da instância perfilada (ex.: “Ele precisava de um emprego, mas não havia *nenhum* emprego disponível”).

Especificação de tipo

Instanciação

As especificações de tipo, por sua vez, representam diferentes níveis de abstração, de modo que uma *hierarquia de tipos* pode conter uma sequência ordenada (ex. COISA > ANIMAL > MAMÍFERO > GATO). Em termos de significado, cada especificação de tipo é esquemática em relação à que lhe segue; além disso, os membros de cada categoria incluem aqueles da próxima como um subconjunto.

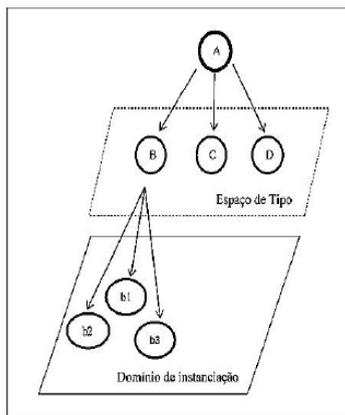
Costumam-se distinguir as relações de tipo/subtipo, por um lado, das relações (sub)tipo/instância, de outro. Para tratar das primeiras, usa-se o termo *elaboração*, enquanto as últimas são designadas como *instanciação*. Langacker (1991) argumenta, entretanto, que a distinção não é absoluta, e que a instanciação pode ser mais bem compreendida como um caso especial de elaboração. Para defender essa posição, o autor analisa expressões nas quais os tipos são construídos como instâncias, como no exemplo a seguir:

(1) O gato é um mamífero que pode viver entre quinze e quarenta anos.

Em (1), o SN “o gato” recebe interpretação genérica, que diz respeito ao animal como tipo – não se trata aqui de um gato particular e espacialmente instanciado. Ao mesmo tempo, entretanto, esse tipo é construído como instanciação de um tipo mais alto (MAMÍFERO).

Para representarmos a estrutura conceptual associada ao uso do SN em (1), consideremos, primeiramente, a situação normal, na qual instanciações constituem o nível mais baixo da hierarquia de tipo:

FIGURA 3 – Espaço físico como domínio de instanciação

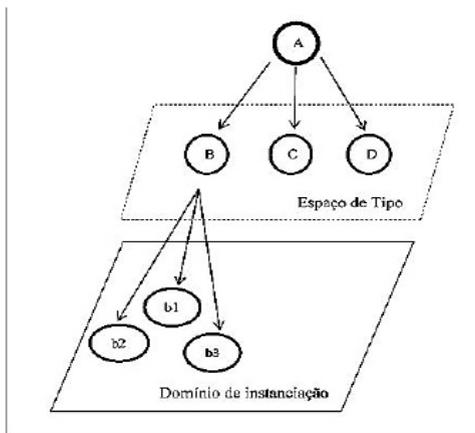


A estrutura A representa um tipo (ex.: MAMÍFERO), enquanto B, C e D são subtipos (ex.: GATO, HUMANO, BALEIA), cada qual com múltiplas instanciações. Por exemplo, as instanciações de B (b1, b2 e b3) podem corresponder a gatos específicos, como ilustram os exemplos a seguir:

- (2) Meu gato sumiu.
- (3) O gato do vizinho é cinza azulado.

A situação é um pouco diferente quando a expressão “o gato” caracteriza um tipo de mamífero (como em (1)). Nesse caso, o domínio pertinente de instanciação não é mais o espaço físico, mas o Espaço de Tipo para MAMÍFERO, como representado a seguir:

FIGURA 4 – Espaço de Tipo como domínio de instanciação



A Figura 4 indica que a especificação de tipo para GATO não é apenas um subtipo de MAMÍFERO, mas instanciações quantificáveis dessa categoria. Nesse caso, pode-se mencionar “dois mamíferos” (ex.: “o gato e o cachorro são dois mamíferos”), do mesmo modo que se pode falar de “dois gatos” em referência a “meu gato” e “o gato do vizinho”.

2.2.2 Grounding

Como descrito anteriormente, considera-se que todo SN perfila uma coisa (*thing*), construindo-a como uma instância quantificada de um determinado tipo e ancorada ao *Ground* – conceito que se refere à situação comunicativa e inclui, portanto, falante, ouvinte, local e momento do ato de fala. Langacker recorre ainda ao termo *grounding predication* (“predicação de ancoragem”) para se referir a predicações que relacionam uma instância perfilada a um determinado ponto de referência.

O falante sempre pode contar com dois pontos de referência naturais, inerentes à situação comunicativa: o *Ground* propriamente dito e uma massa de referência (R_T), que inclui todas as instâncias no âmbito do discurso. Por exemplo, para o nome “menino”, que designa uma entidade discreta, R_T é equivalente ao conjunto de todas as instâncias.

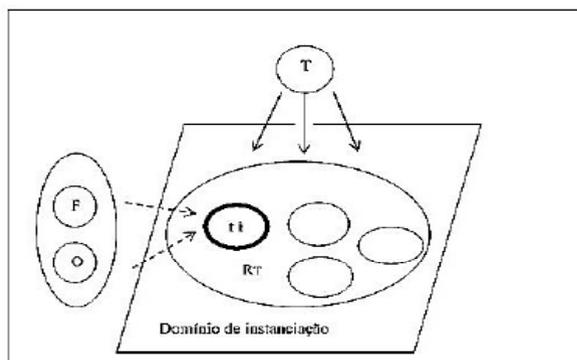
O objetivo do falante ao usar um SN é mencionar uma determinada instância de um tipo, de modo que o ouvinte possa determinar a referência pretendida. Por exemplo, para um determinado tipo T , existe normalmente um conjunto aberto de instâncias (t_1, t_2, t_3, \dots). Além disso, para se determinar o valor definido (ou indefinido) da instância perfilada, o ponto de referência externo inclui falante e ouvinte, que se veem sob o desafio de dirigir a atenção para a mesma instância, sendo responsáveis pela conceptualização que constitui o polo semântico do SN.

Assim, falante (F) e ouvinte (O), que formam conjuntamente o *Ground*, enfrentam a tarefa de coordenar suas referências mentais a uma determinada instância t_1 do tipo T , destacada da massa de referência R_T . Tanto o *Ground* quanto R_T estão disponíveis como pontos de referência para esse propósito.

Tendo em vista que os participantes do evento de fala estabelecem contato com a entidade concebida e a relacionam ao seu próprio conhecimento das instâncias, as predicções de ancoragem são consideradas epistêmicas. Ao usar o *Ground* como ponto de referência, o indivíduo estabelece *contato mental* com t_1 e, portanto, t_1 é destacado para tomada de consciência no estado mental corrente do falante.

A Figura 5 representa os fenômenos de instanciação e contato mental, em que há coordenação plena da referência:

FIGURA 5 – Instanciação e contato mental



Com base na Figura 5, conclui-se que uma predicção epistêmica sempre perfila a entidade ancorada, e não a relação de ancoragem em si mesma (embora essa relação seja parte essencial de sua base e, portanto, de seu significado). Como já mencionamos, a entidade perfilada para as

expressões nominais é uma *coisa*: uma instância quantificada e única do tipo T e retirada da massa de referência RT. A predicação de ancoragem evoca os interactantes como pontos de referência e faz algum tipo de especificação referente ao estabelecimento de contato mental com t_1 por parte deles.

2.2.3 Definitude e indefinitude

Neste trabalho, interessam-nos, em especial, as predicções que enfocam grau de definitude e tomam o *Ground* como ponto de referência primário.

Para caracterizar o artigo definido, Langacker (1991) recorre à noção de *current discourse space* (Espaço Discursivo Corrente, ou apenas EDC), que contém os elementos e relações construídos como compartilhados por falante e ouvinte. Esses elementos e relações constituem a base para a comunicação em um determinado momento do fluxo discursivo, e representam aquilo que é concebido como imediatamente disponível para os interlocutores, figurando diretamente na consciência ou sendo prontamente evocado por associação ou inferência. Outra noção básica é a de *contato mental*, definida como um processo em que uma entidade é isolada (*singled out*) para tomada de consciência individual. Com base nas noções de EDC e contato mental, Langacker (1991, p. 98) propõe a seguinte caracterização do artigo definido:

O uso do artigo definido como descrição de tipo T em um sintagma nominal implica que: (1) a instância designada t_1 de T é única e máxima em relação ao EDC; (2) o Falante tem contato mental com t_1 ; e (3) o Ouvinte pode ter contato mental com t_1 ou o próprio SN é suficiente para estabelecer esse contato⁵.

Segundo essa definição, o artigo em si mesmo perfila t_1 , e seu uso implica que o SN é suficiente para colocar o ouvinte em contato mental com uma instância unicamente identificada, sem depender de outras informações disponíveis na cláusula que o contém.

⁵“The use of the definite article with type description T in a nominal implies that (1) the designated instance t_1 of T is unique and maximal in relation to the current discourse space; (2) S has mental contact with t_1 ; and (3) either H has mental contact with t_1 or the nominal alone is sufficient to establish it.”

A especificação de que t_1 é único no EDC pode ser ilustrada com os seguintes exemplos:

- (4) a. Douglas comeu uma maçã e um figo. A maçã estava deliciosa.
 b. *Douglas comeu uma maçã verde e uma maçã vermelha. A maçã estava deliciosa.

Em (4^a), o EDC criado pela primeira cláusula contém apenas uma instância de MAÇÃ. Sendo assim, a referência pretendida é inequívoca. Já em (4b), o EDC contém duas instâncias de MAÇÃ; dessa forma, t_1 não é único e o artigo definido “o” é inapropriado.

A especificação de que t_1 é máximo estabelece que t_1 só pode ser identificado com a instância mais inclusiva no espaço discursivo, como ilustram os seguintes exemplos:

- (5) a. José tem três carros. Os carros são esportivos.
 b. José comprou vinte litros de água. A água foi usada para molhar as plantas.

Nos exemplos, as primeiras cláusulas introduzem uma instância de CARRO, com a cardinalidade de três (ex. (5a)) e uma instância de ÁGUA, com um volume de vinte litros (ex. (5b)). Trata-se de instâncias máximas no espaço discursivo, às quais “os carros” e “a água” se referem com propriedade.

Assim como os definidos, os SNs com artigo indefinido também tomam o *Ground* como ponto de referência primário. Estes últimos, contudo, apresentam uma diferença crucial em relação aos primeiros, por serem insuficientes para colocar o ouvinte em contato mental com uma instância única de T. Langacker retoma aqui a proposta de Hawkins (1978), para quem o artigo indefinido contrasta com o definido em relação à unicidade: o primeiro implica que o SN em si mesmo não é suficiente para colocar o ouvinte em contato mental com uma instância unicamente determinada da categoria. Assim, se há dois mecânicos consertando um carro, e vários parafusos disponíveis, é possível que um diga para o outro “Poderia me passar um parafuso?” – mas dificilmente diria “Poderia me passar o parafuso?”. Isso porque há múltiplas instâncias de parafuso no EDC, definido neste caso pelas circunstâncias físicas imediatas.

Com relação aos sintagmas indefinidos, Langacker (1991) assinala que há uma clara divisão de função entre os que apresentam artigo indefinido e os que ocorrem com zero: o artigo indefinido ocorre apenas com nomes singulares contáveis (“Ele comprou um livro”), enquanto o determinante zero ocorre com nomes contáveis e não-contáveis (“Alice bebeu (Ø) leite”; “Ø Ratos entraram na loja”). Além disso, apenas o determinante zero admite um *construal* plenamente genérico (“Ø Cachorros são mamíferos”).⁶

De especial relevância são os casos que envolvem a noção de instância arbitrária. Uma delas é a chamada leitura opaca, que constitui uma das leituras possíveis da sentença *Maria quer comprar um vestido*. Aqui, o referente VESTIDO não é acessado no espaço de ‘realidade’, mas no espaço de desejo de Maria – trata-se, portanto, de um vestido não-específico.⁷ Outro caso diz respeito ao valor do artigo indefinido em construções de predicado nominal (ex.: “Alice é uma ladra”). Por fim, há o uso que envolve sentenças genéricas, como “Uma criança sempre tem sonhos”. O que todos esses usos têm em comum é o fato de poderem ser considerados casos de não-especificidade, nos quais uma determinada instância t_1 de T é evocada para um propósito particular limitado, e só se sustenta em contextos desse tipo. É importante destacar que essas sentenças não implicam que falante e ouvinte tenham contato mental pré-existente com a instância designada pelo sintagma indefinido. A instância, na verdade, é evocada com vistas a se estabelecer uma afirmação genérica e como tal é concebida como uma instância representativa da categoria, e não como uma instância particular, conhecida em bases independentes. Assim, tanto “uma criança” quanto “sonhos” podem suscitar leituras não-específicas, de forma semelhante a “um vestido” em interpretação opaca na sentença “Maria quer comprar um vestido”.

Há um caso de sintagma singular, entretanto, que não foi discutido por Langacker. Trata-se do sintagma singular genérico com determinante zero, cuja ocorrência não é gramatical em inglês (“*Cat is lazy”). Esse

⁶ Para uma discussão mais detalhada desses casos, remetemos a Langacker (1991, p. 89-95 e p. 103-107).

⁷ A sentença também admite a leitura transparente em que VESTIDO é acessado na Base (espaço de ‘realidade’). Essa leitura, por ser específica, não se inclui nos casos de instância arbitrária discutidos por Langacker (1991).

tipo de estrutura, contudo, ocorre em português (“Gato é preguiçoso”), e se inclui no foco de investigação deste trabalho. Na próxima seção, apresentaremos uma proposta teórica que busca ampliar a proposta langackeriana, a fim de contrastar os usos genéricos dos sintagmas singulares com artigo definido e com determinante zero no PB.

3 SNs singulares genéricos no PB: uma abordagem cognitivista

Nesta seção, propomos uma análise cognitivista para dois tipos de SN singulares genéricos do PB: o padrão com determinante zero, como em (6^a) e (7^a), e o padrão com artigo definido, como em (6b) e (7b).

(6) a. Gato é muito voluntarioso.

b. O gato é muito voluntarioso.

(7) a. Brasileiro gosta de feijoada.

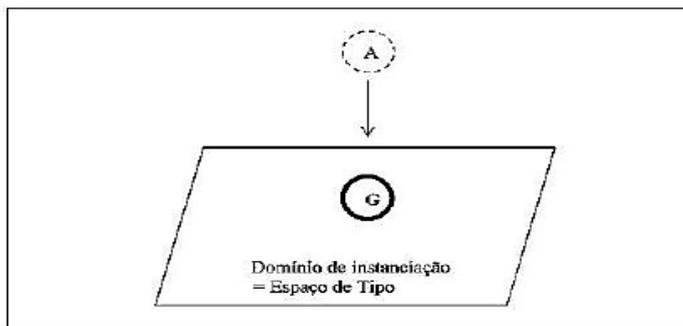
b. O brasileiro gosta de feijoada.

Embora nosso objetivo último seja o de explorar as *diferenças* entre essas estruturas, vale a pena iniciar a análise assinalando as semelhanças entre elas. Em primeiro lugar, ambos os padrões designam *tipos* (em oposição a instâncias particulares), o que significa que tanto o esquema ARTIGO DEFINIDO + SN quanto o esquema Ø + SN perfilam uma categoria inserida em um Espaço de Tipo, e não uma entidade particular inserida no espaço físico. Assim, quando alguém enuncia (6a) ou (6b), está fazendo referência a uma categoria (GATO) que tem existência em um espaço abstrato (digamos, o espaço abstrato das espécies animais, ou dos mamíferos, ou dos felinos), e não a um ser particular (digamos, o Tom) que tem existência no espaço físico (ainda que, no caso do Tom, não se trate do espaço físico do mundo real). Da mesma maneira, quando alguém enuncia (7a) ou (7b), está se referindo à categoria genérica BRASILEIRO, que existe no espaço abstrato das nacionalidades, e não a qualquer brasileiro particular que habite o mundo físico.

Nos termos da GC, essa proposta pode ser traduzida assim: quando se trata de SNs singulares genéricos, o domínio da instanciação do referente do núcleo do SN corresponde a um Espaço de Tipo – isto

é, a um domínio conceptual abstrato. Tomando como exemplos os SNs sujeitos de (6), essa ideia pode ser representada assim:

FIGURA 6 – Representação conceptual parcial das expressões genéricas “Gato” e “O gato”



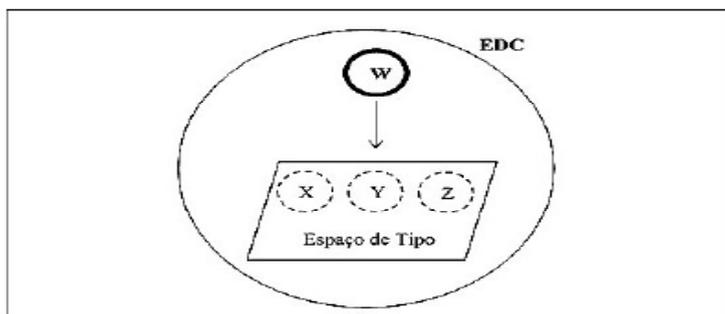
A Figura 6 ilustra uma hierarquia de tipo na qual a letra A representa o tipo ANIMAL e a letra G representa o tipo GATO. Usamos a linha forte no círculo ao redor da letra G para indicar que o tipo GATO é conceptualmente *perfilado* pelas expressões “Gato” e “O gato”; pela mesma lógica, usamos a linha pontilhada no círculo ao redor da letra A para indicar que o tipo ANIMAL corresponde à porção não-perfilada da predicação (isto é, ao que Langacker chamada de *base*).

Se a representação acima captura as similaridades entre os dois padrões de SN genérico analisados aqui, é preciso agora investigar as diferenças entre eles. Para isso, dividiremos nossa hipótese em duas partes: (i) uma hipótese referente às representações conceptuais compartilhadas pelos interlocutores – e, portanto, disponíveis no EDC – no momento *imediatamente anterior* ao emprego de cada um dos padrões e (ii) uma hipótese referente às representações conceptuais que são acionadas – e, portanto, passam a estar disponíveis no EDC – em função do próprio emprego de cada um dos padrões de SN genérico. No primeiro caso, trata-se de investigar o *contexto discursivo que licencia* o emprego de uma ou outra estrutura; no segundo, trata-se de descrever o *cenário conceptual diretamente predicado* pelas estruturas investigadas.

Em relação ao primeiro ponto, sugerimos que o padrão definido é licenciado nos contextos em que o domínio conceptual de instanciação do referente do núcleo do SN já se encontra previamente disponível.

Nos termos da GC, isso significa que o Espaço do Tipo que abriga o tipo designado pelo SN deve estar previamente acessível no Espaço Discursivo Corrente. Essa situação pode ser representada assim⁸:

FIGURA 7 – Representação do contexto discursivo que licencia o padrão com artigo

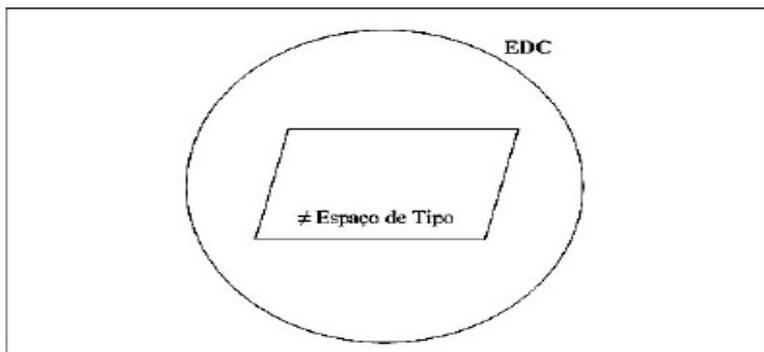


A Figura 7 busca capturar o conhecimento compartilhado entre os participantes da situação comunicativa no momento anterior à enunciação de um SN singular genérico com artigo definido. A presença do tipo W representa o fato de que, nesse ponto da interação, um domínio conceptual abstrato deve estar acessível ao Falante e ao Ouvinte. Por definição, o acesso mental a essa categoria faculta a conceptualização de um Espaço de Tipo, bem como dos tipos inseridos nele. Indicados no diagrama acima pelas variáveis X, Y e Z, esses tipos estão representados por linhas pontilhadas para marcar o fato de que, nesse momento, eles não estão perfilados, dado que ainda não foram mencionados de forma explícita. Em outras palavras, sua conceptualização decorre tão somente do conhecimento de mundo compartilhado pelos interactantes. A Figura 7 representa, portanto, uma situação na qual uma categoria mais geral (como ANIMAL, por exemplo) já está acessível, levando à conceptualização de um Espaço de Tipo (por exemplo, o espaço conceptual abstrato das espécies animais) e à “visualização” implícita (ou “offstage”, nos termos de Langacker) de tipos pertencentes a esse espaço (por exemplo, GATO, CACHORRO e MORCEGO).

⁸ Como o Espaço Discursivo Corrente é tratado por Langacker (1991, p. 97) como um tipo de espaço mental, nos termos do modelo de Fauconnier (1994), iremos representá-lo como um círculo, como tradicionalmente se faz na Teoria dos Espaços Mentais.

No que diz respeito ao padrão com determinante zero, sugerimos que o contexto discursivo que o licencia é menos específico – fundamentalmente, ele é caracterizado pela *ausência* de uma categoria geral (como ANIMAL, ou NACIONALIDADE, ou qualquer outra) no Espaço Discursivo Corrente. Em outras palavras, sugerimos que a formulação sem artigo será preferida quando o Espaço de Tipo – que, como vimos, funciona como domínio de instanciação do referente do núcleo do SN – *não* estiver disponível no EDC. Essa ideia pode ser representada assim:

FIGURA 8 – Representação do contexto discursivo que licencia o padrão com determinante zero



A Figura 8 representa uma situação na qual o domínio conceptual abstrato que sustenta a instanciação do tipo designado pelo SN genérico *não* se encontra no EDC, por não ter sido ativado no contexto discursivo. Por essa razão, também não estão representados no diagrama acima tipos abstratos (como aqueles indicados pelas variáveis X, Y e Z na Figura 7). Em termos práticos, isso significa que uma diferença crucial entre as estruturas com artigo e com determinante zero diz respeito à disponibilidade prévia da entidade a ser perfilada pelo SN: estamos sugerindo que, no primeiro caso, essa entidade está previamente disponível (ainda que “offstage”) ao passo que, no segundo caso, ela não comparece no EDC.

Isso nos leva à segunda parte da nossa hipótese – aquela que diz respeito à representação conceptual evocada pelo próprio SN singular genérico. Se a alternativa com artigo é empregada em uma situação interacional como aquela representada na Figura 7, isso significa que seu papel é tão somente o de *perflar* uma entidade que já se encontra

previamente disponível no EDC (em detrimento de outras entidades que igualmente povoam o Espaço de Tipo). Em contraste, se a alternativa sem artigo é empregada no contexto representado na Figura 8, isso significa que seu papel será o de inserir uma nova entidade na EDC. Essa diferença fica evidenciada pelo contraste entre as figuras 9 e 10:

FIGURA 9 – Representação conceptual associada ao SN singular genérico com artigo definido

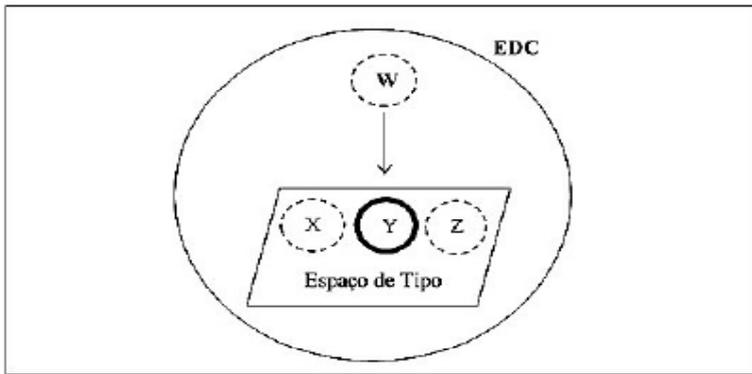
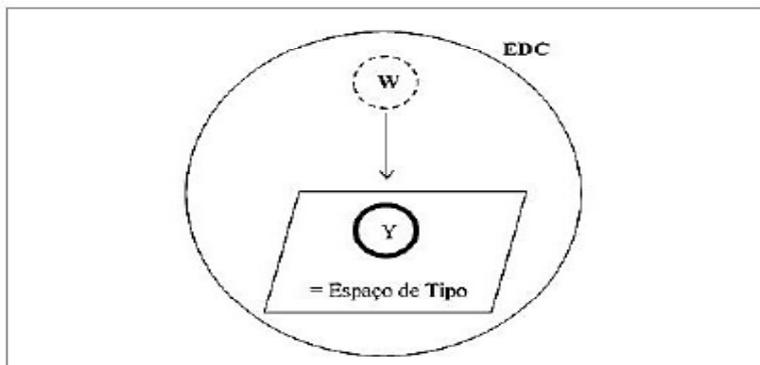


FIGURA 10 – Representação conceptual associada ao SN singular genérico com determinante zero



A Figura 9 representa o fato de que o papel do SN singular genérico com artigo é o de perfilar (conforme indicado pela linha mais forte ao redor do Y) uma entidade já presente do Espaço de Tipo abstrato que serve como domínio de instanciação para o referente do seu núcleo (conforme indicado na Figura 7). A Figura ajuda ainda a explicitar visualmente a ideia de que o padrão com artigo *seleciona* uma entidade conceptual (aquela representada por Y) em meio a um conjunto de tipos não-perfilados (aqueles representados por X e Z).

A Figura 10, por seu turno, representa a ideia de que o SN singular genérico com determinante zero estabelece no Espaço Discursivo Corrente um novo Espaço de Tipo, ao mesmo tempo em que perfila um tipo inserido nesse espaço (aqui representado pela variável Y). A comparação entre as figuras 9 e 10 deixa evidente que a diferença entre os dois padrões de SN singular genérico diz respeito à parte não-perfilada da predicação (isto é, ao que Langacker chama de base): crucialmente, apenas a alternativa com artigo faculta a “visualização” de tipos não-perfilados presentes no domínio de instanciação, para além daquele diretamente designado pelo SN. As diferenças entre as duas estruturas estão sintetizadas na tabela abaixo:

TABELA 1 – Diferenças entre os padrões de SN singular genérico com e sem artigo definido

	REPRESENTAÇÃO CONCEPTUAL ASSOCIADA AO CONTEXTO PRÉVIO	REPRESENTAÇÃO CONCEPTUAL ASSOCIADA AO SN SINGULAR GENÉRICO
ARTIGO DEFINIDO + NOME	Espaço de Tipo disponível no EDC; tipos disponíveis no EDC mas não perfilados	Perfilamento de um dos tipos inseridos no Espaço de Tipo previamente disponível no EDC
DETERMINANTE ZERO + NOME	Espaço de Tipo e tipos não disponíveis no EDC	Estabelecimento de um novo Espaço de Tipo e de um novo tipo no EDC

Conforme observamos na Introdução deste artigo, a Gramática Cognitiva tem recebido algumas críticas em função da alegada ausência, ou pelo menos escassez, de suporte experimental para seus constructos

teóricos (HOLLMANN, 2013; BROCCIAS; HOLLMANN, 2007) – nos termos de Hart (2014, p. 16), “a realidade psicológica da Gramática Cognitiva é, neste momento, plausível mas ainda não demonstrada”⁹. Dessa maneira, análises calcadas na GC podem ser consideradas hipóteses produtivas e instigantes, mas dificilmente poderiam prescindir de algum tipo de verificação empírica.

Isso vale, naturalmente, para a análise desenvolvida nesta seção. Por um lado, é possível que alguns a considerem produtiva e instigante – ela pode ajudar a entender, por exemplo, por que expressões com SN singular genérico com e sem artigo *parecem* sinônimas (porque perfilam um mesmo tipo de entidade, em um mesmo tipo de Domínio de instanciação), ao mesmo tempo em que esclarece por que elas, na verdade, *não são* sinônimas (porque se distinguem quanto à porção não-perfilada – isto é, implícita – da predicação). Mas até que ponto esse tipo de sugestão é confiável? Que evidências nós temos de que a percepção de uma sentença como (6b) dispara a conceptualização implícita de outros animais (digamos, CACHORRO, MORCEGO, etc.), ao passo que o mesmo efeito não se verifica com uma sentença do tipo (6^a)? É para essa questão que nos voltamos na próxima seção.

4 Testando a hipótese

A fim de verificar a realidade psicológica da análise desenvolvida na seção anterior, foi realizado um teste offline de julgamento de aceitabilidade. Nesta seção, descrevemos o experimento (item 4.1) e, na sequência, apresentamos e discutimos seus resultados (item 4.2).

4.1 O experimento: visão geral

Para o experimento de julgamento de aceitabilidade, foi utilizada escala Likert de cinco pontos. Como variável dependente, portanto, temos o grau de aceitabilidade do estímulo, que poderia variar entre os níveis 1 (completamente inaceitável) e 5 (completamente aceitável). A única variável independente considerada foi a presença ou ausência de contraste. Assim, os estímulos críticos foram divididos em somente duas condições – sem contraste e com contraste –, conforme os exemplos abaixo:

⁹“The psychological reality of Cognitive Grammar is, at this point in time, plausible but not yet proven”.

(8) A criança gosta de doces e guloseimas.

(9) A criança gosta de doces e guloseimas. Já o adulto tende a ter um paladar mais apurado.

Os estímulos da condição sem contraste, como (8), predicam sobre uma única categoria (no exemplo acima, CRIANÇA). Diferentemente, os estímulos da condição com contraste, como (9), estabelecem uma oposição entre duas categorias distintas (no exemplo acima, CRIANÇA e ADULTO).¹⁰ Neste experimento, prevemos que o grau de aceitabilidade do estímulo estará relacionado à presença ou ausência de contraste, de maneira tal que *os estímulos com contraste serão tidos como mais aceitáveis que os estímulos sem contraste* (mas ver a seção 4.1.3 para uma outra formulação da previsão experimental).

Para compreender a lógica por trás do experimento, é importante lembrar a hipótese desenvolvida na seção 3. Resumidamente, defendemos que o SN singular genérico com artigo definido (i) pressupõe a disponibilidade prévia do Espaço de Tipo que funcionará como domínio de instanciação do referente do seu núcleo e (ii) faculta a conceptualização implícita de tipos não-perfilados. Ao mesmo tempo, e inversamente, sugerimos que o SN singular genérico com determinante zero (i) não pressupõe a disponibilidade prévia do Espaço de Tipo que funcionará como Domínio de instanciação e (ii) não faculta a conceptualização implícita de tipos não-perfilados.

Ao manipular a presença e ausência de sequências contrastivas nos estímulos experimentais, foi possível criar estímulos mais e menos compatíveis com a descrição sugerida para o padrão de SN genérico com artigo. Especificamente, a inserção de uma sequência adicional nos estímulos com contraste favorece a inferência de que o domínio conceptual do ET corresponde ao próprio tópico do discurso. Por exemplo, a segunda sentença de (9) favorece a interpretação de que as sentenças presentes no estímulo são parte de uma sequência discursiva mais ampla na qual o tópico em pauta são as diferenças entre as etapas da vida humana. Tecnicamente, realizar esse tipo de inferência equivale a assumir que (i) o Espaço de Tipo FASES DA VIDA HUMANA estava previamente disponível para os interlocutores hipotéticos da situação comunicativa em que a

¹⁰ Para detalhes sobre a composição dos estímulos, ver seção 4.1.2.

sentença teria sido enunciada e (ii) a função realizada pelo SN genérico foi a de perfilar um tipo particular (por exemplo, o tipo CRIANÇA) em meio a outros tipos conceitualmente disponíveis (por exemplo, ADOLESCENTE e ADULTO). Assumindo-se que os sujeitos experimentais de fato realizarão essa inferência, essa proposta produz a previsão de que os estímulos com contraste deverão ser considerados naturais e bem-sucedidos.¹¹

Nos estímulos da condição sem contraste, por outro lado, a ausência de uma sequência contrastiva torna menos evidente ou imediata a interpretação de que o domínio conceptual do ET corresponde ao próprio tópico do discurso.¹² Em outras palavras, assumimos que, em uma sentença como (8), os sujeitos experimentais terão menos subsídios – em comparação com a condição com contraste – para interpretar o enunciado como parte de uma sequência discursiva mais ampla sobre as FASES DA VIDA HUMANA. Tecnicamente, isso significa que eles estarão menos propensos a assumir que (i) o ET FASES DA VIDA HUMANA estava previamente disponível para os interlocutores e (ii) o SN genérico teria cumprido tão somente a função de perfilar um tipo particular em meio a um conjunto de tipos disponíveis. Caso a nossa hipótese esteja correta, a ausência (ou menor disponibilidade) dessa interpretação deverá conduzir a um grau de aceitabilidade relativamente mais reduzido para os estímulos sem contraste.

Em resumo, as condições experimentais procuram manipular o acionamento de inferências contextuais por parte dos sujeitos do experimento: enquanto a inserção de uma sequência contrastiva deverá favorecer a evocação do tipo de representação conceptual associado ao SN singular genérico com artigo, a supressão dessa mesma sequência deverá produzir o efeito inverso. Como, no entanto, todos os estímulos críticos contêm SNs genéricos *com* artigos, espera-se que essa diferença resulte em uma diferença significativa na aceitabilidade dos estímulos pertencentes a cada uma das condições.

¹¹ Naturalmente, o desenho do experimento e a composição dos estímulos pressupõem a ideia de que ouvintes/leitores são capazes de fazer inferências sobre o contexto em que os enunciados ouvidos/lidos teriam sido utilizados (WESTERA; BRASOVEANU, 2014; SAVIC, 2014).

¹² O que não significa que essa interpretação seja impossível – voltaremos a esse ponto na seção 4.2.

4.1.1 Desenho experimental, estímulos e participantes

Participaram voluntariamente do experimento 30 sujeitos (21 mulheres e 9 homens), com idades entre 18 e 57 anos, todos estudantes do 1º período da Faculdade de Letras da UFRJ e falantes nativos do PB. Em troca da participação, os voluntários receberam horas de Atividades Acadêmico-científico-culturais (AACCs).

Ao todo, foram elaborados 16 estímulos críticos (oito para cada condição) e 16 distratores. Por meio de delineamento em quadrado latino, os estímulos críticos foram divididos em dois scripts (aqui referidos como scripts A e B) a fim de evitar a repetição de material lexical entre as duas condições experimentais. Como se optou por um desenho intra-sujeitos, cada participante julgou quatro estímulos críticos com contraste e quatro estímulos críticos sem contraste (além de dezesseis distratores), totalizando 240 observações.

Para a criação dos estímulos, foram definidos previamente sete domínios semânticos: *alimento*, *bebida*, *fases da vida humana*, *meio de transporte*, *equipamento eletrônico*, *esporte* e *roupa*. Para o domínio *alimento* foram estabelecidos dois pares de classes contrastantes (*chocolate vs. Maçã* e *batata frita vs. Fruta*); além disso, para cada um dos seis domínios restantes, foi estabelecido um par de classes contrastantes (respectivamente, *leite vs. Chá*; *criança vs. Adulto*; *carro vs. Metrô*; *notebook vs. Tablet*; *futebol vs. Golfe*; *casaco vs. Regata*). Por fim, para cada um desses oito pares opostos, foram criadas duas sentenças – uma com contraste e uma sem contraste –, resultando em 16 estímulos críticos. Todos os estímulos podem ser vistos no anexo 1.

Os estímulos críticos foram construídos segundo o modelo das sentenças (6) e (7), repetidas abaixo, por conveniência, como (8) e (9):

(8) A criança gosta de doces e guloseimas.

(9) A criança gosta de doces e guloseimas. Já o adulto tende a ter um paladar mais apurado.

Como se vê, os estímulos críticos sempre exibem artigo definido (não há, portanto, estímulos com determinante zero).¹³ Em seqüências pertencentes à condição sem contraste, como (8), temos uma sentença simples e apenas um SN singular genérico, sempre na posição de sujeito. Por seu turno, seqüências pertencentes à condição com contraste, como (9), apresentam duas sentenças simples, com um SN singular genérico em cada uma delas. Tomadas conjuntamente, essas duas sentenças descrevem uma situação de contraste: o referente do sujeito da primeira sentença (codificado como SN singular genérico com artigo definido) é contrastado, com base em algum atributo, ao referente do sujeito da segunda sentença (igualmente codificado como SN singular genérico com artigo definido). Em (9), por exemplo, o atributo relevante diz respeito à sofisticação do paladar: afirma-se, assim, que membros da classe CRIANÇA apresentam paladar menos sofisticado que membros da classe ADULTO.

4.1.2 Materiais e procedimentos

O experimento foi montado no software PsychoPy, versão 1.84.2, e aplicado em um notebook Positivo Duo ZX3020, com tela de 10.1 polegadas. Inicialmente, todos os voluntários receberam explicações orais sobre a tarefa experimental e sobre a dinâmica do experimento. Em seguida, cada sujeito avaliou, na presença do pesquisador, cinco estímulos de treinamento. Depois disso, os participantes foram deixados por conta própria.

Em um primeiro momento, os sujeitos tinham acesso a uma tela inicial com instruções escritas (anexo 2), que poderiam ler quantas vezes quisessem e por quanto tempo julgassem necessário. A fim de assegurar a interpretação genérica (e não específica) dos estímulos críticos, esse texto inicial solicitava ao participante que imaginasse uma situação hipotética na qual um pai estivesse transmitindo ao seu filho “uma série de ensinamentos sobre diversos assuntos e seres”. Em seguida, o mesmo texto solicitava que os participantes julgassem a aceitabilidade das frases produzidas por esse pai hipotético, a partir de uma escala com cinco níveis:¹⁴

¹³ Conforme a proposta desenvolvida na seção 3, o padrão com determinante zero é não-marcado em relação ao conteúdo inserido no EDC antes do momento da enunciação da sentença. Por essa razão, a hipótese teórica não prevê que esse padrão seja necessariamente incompatível com a existência de seqüências contrastivas. É isso que justifica a opção por incluir no experimento apenas estímulos com artigo definido.

¹⁴ O texto com as instruções iniciais pode ser visto no anexo 2.

1. Totalmente inaceitável
2. Pouco aceitável
3. Medianamente aceitável
4. Muito aceitável
5. Totalmente aceitável

Após ter lido e compreendido as instruções iniciais, o participante deveria pressionar a barra de espaço para que se iniciasse a apresentação dos estímulos. As 24 sequências textuais a serem julgadas por cada sujeito foram organizadas em ordem pseudoaleatória, de maneira que sequências críticas não fossem exibidas em estímulos consecutivos. Cada estímulo ficou exposto por 4000ms, sendo então seguido da tela mostrada na Figura 11:

FIGURA 11 – Tela da pergunta referente ao julgamento em Escala Likert



Essa tela, por seu turno, permanecia visível por tempo indeterminado, até que os participantes indicassem, por meio dos números do teclado, o grau de aceitabilidade da sequência exposta na tela imediatamente anterior. Essa resposta levava automaticamente à exibição do estímulo seguinte.

4.1.3 Análise estatística e previsões experimentais

Para a análise estatística, foi realizado um teste qui-quadrado de homogeneidade por meio do software Action Stat, versão 3.1.43.724.694. Optou-se pela técnica do qui-quadrado, em detrimento da mais tradicional Análise de Variância, porque existe questionamento na literatura quanto à possibilidade de se tratar os dados derivados de escala Likert como

dados numéricos (JAMIESON, 2004). Aqui, portanto, assumimos a opção metodologicamente mais conservadora de tratar o grau de aceitabilidade como variável categórica.

Especificamente, o qui-quadrado de homogeneidade avalia se duas ou mais populações apresentam distribuição equivalente de uma dada variável categórica. No nosso caso, as populações correspondem às duas condições experimentais (com e sem contraste), ao passo que a variável categórica corresponde ao próprio grau de aceitabilidade atribuído a cada estímulo pelos participantes do experimento. Dessa maneira, o qui-quadrado de homogeneidade permite avaliar se a proporção de sequências com grau de aceitabilidade 1, 2, 3, 4 e 5 é equivalente nas duas condições.

Mais acima, apresentamos informalmente nossa previsão experimental da seguinte maneira: os estímulos da condição com contraste serão considerados mais aceitáveis do que os estímulos da condição sem contraste. Como não trabalharemos com médias numéricas, e sim com contagem de frequência dos graus de aceitabilidade, formularemos essa previsão nos seguintes termos: *a proporção de sentenças totalmente inaceitáveis (grau 1), pouco aceitáveis (grau 2), medianamente aceitáveis (grau 3), muito aceitáveis (grau 4) e totalmente aceitáveis (grau 5) não será equivalente nas duas condições experimentais*. Se estivermos corretos, esse resultado revelará a existência de uma correlação estatística entre a presença ou ausência de contraste, de um lado, e o grau de aceitabilidade de sentenças com SNs singulares genéricos com artigo, de outro.

4.2 Resultados e discussão

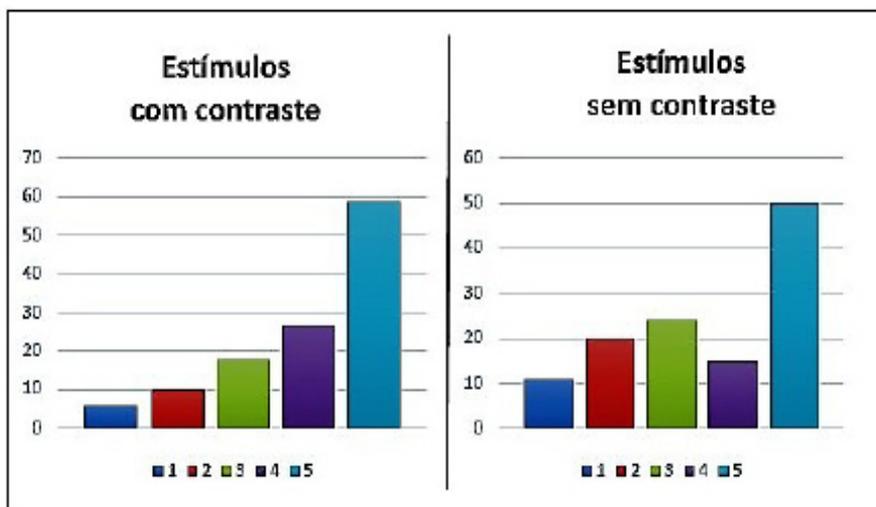
Conforme previsto pela hipótese teórica desenvolvida na seção 3, existe diferença significativa na distribuição dos graus de aceitabilidade dos estímulos entre as duas condições experimentais ($\chi^2(4) = 9.832, p = .043$), com um efeito de intensidade moderada (V de Cramer = .2).¹⁵ A tabela 2 e o gráfico 1 sintetizam a distribuição dos graus de aceitabilidade em cada condição:

¹⁵ Tipicamente, assume-se que valores de V de Cramer iguais ou maiores que .1 e menores que .2 indicam efeito fraco, valores iguais ou maiores que .2 e menores que .4 indicam efeito moderado, valores iguais ou maiores que .4 e menores que .6 indicam efeito forte e valores acima de .6 indicam efeito muito forte (REA; PARKER, 1992).

TABELA 2 – Grau de aceitabilidade dos estímulos críticos

Grau de aceitabilidade	COM CONTRASTE		SEM CONTRASTE		TOTAL
	Número de respostas	Valor percentual	Número de respostas	Valor percentual	
1	6	5%	11	9.16%	17
2	10	8.33%	20	16.66%	30
3	18	15%	24	20%	42
4	27	22.5%	15	12.5%	42
5	59	49.16%	50	41.66%	109
Total	120		120		240

GRÁFICO 1 – Distribuição dos dados nas condições com e sem contraste



O padrão que emerge dos dados acima é o seguinte: o índice de sentenças totalmente inaceitáveis (grau 1), pouco aceitáveis (grau 2) e medianamente aceitáveis (grau 3) é mais elevado na condição sem contraste, ao passo que o número de sentenças muito aceitáveis (grau 4) e totalmente aceitáveis (grau 5) é maior na condição com contraste. Especificamente, as sequências julgadas completamente aceitáveis perfazem aproximadamente 49.16% do total de estímulos com contraste e

41.66% dos estímulos sem contraste; de maneira semelhante, 22.5% dos estímulos contrastivos receberam grau de aceitabilidade 4, ao passo que o mesmo só ocorreu com 12.5% dos estímulos não-contrastivos. Ao mesmo tempo, quando se trata dos estímulos com menor nível de aceitabilidade (graus 1, 2 e 3), a situação se inverte: as proporções são sistematicamente mais elevadas na condição sem contraste (respectivamente, 9.16%, 16.66% e 20%) do que na condição com contraste (respectivamente, 5%, 8.33% e 15%).¹⁶

Os resultados sugerem, portanto, que a presença de um segundo SN genérico, com função contrastiva, na mesma sequência textual eleva significativamente – ainda que moderadamente – a aceitabilidade de enunciados que contêm SN genérico com artigo definido (e, naturalmente, a ausência do segundo SN genérico produz o efeito inverso). Isto é, o falante está mais propenso a aceitar uma sentença descontextualizada com SN genérico introduzido por artigo definido quando ela é seguida por uma sentença adicional capaz de construir uma relação opositiva.

Isso é precisamente o que se poderia esperar com base na proposta teórica desenvolvida na seção 3. Conforme argumentamos, a sequência contrastiva convida o leitor a evocar o tipo de esquema conceptual associado ao padrão sintático de SN genérico com artigo definido. Nesse esquema, vale lembrar, (i) o Espaço de Tipo que funciona como domínio de instanciação do referente do núcleo do SN se encontra previamente disponível no Espaço Discursivo Corrente e (ii) esse mesmo Espaço de Tipo é ocupado não apenas pelo referente do núcleo do SN, mas também por um conjunto de tipos conceptualmente pressupostos (isto é, “offstage”, ou não-perfilados). Se, diante de um estímulo linguístico, o sujeito é capaz de evocar mentalmente esse cenário (conforme representado nas Figuras 7 e 9), é de esperar que ele perceba o enunciado como natural e aceitável. Ao mesmo tempo, caso o indivíduo *falhe* em construir essa representação (situação mais provável na *ausência* de uma sequência contrastiva), é esperado que ele sinta algum tipo de estranhamento diante de um enunciado com SN genérico definido. Dessa forma, a diferença

¹⁶ Um dos pareceristas argumenta que a diferença entre a proporção de notas 5 para estímulos com e sem contraste é “mínima”. O dado crucial, no entanto, é que a diferença na distribuição dos graus de aceitabilidade nas duas condições é *estatisticamente significativa* e evidencia *efeito moderado*. Estes são, portanto, os dois fatos empíricos a serem explicados. Nesta seção, propomos uma interpretação para eles.

na distribuição dos graus de aceitabilidade entre as duas condições é interpretada aqui como efeito da maior ou menor probabilidade de se evocar o tipo de representação conceptual associado à variante definida do SN singular genérico.

Ao mesmo tempo, é importante lembrar que este experimento foi capaz de detectar um efeito apenas *moderado* da presença do segundo SN sobre a distribuição dos graus de aceitabilidade dos estímulos (V de Cramer = .2). A nosso ver, isso pode ser explicado, ao menos em parte, pela assunção de que leitores/ouvintes jamais interpretam enunciados em um vácuo discursivo e informacional. Isso significa que, mesmo em casos nos quais nenhum contexto discursivo é explicitamente fornecido, eles tenderão a evocar algum tipo de entorno contextual para a sequência que está sendo interpretada¹⁷. Dessa forma, ao se deparar com uma sentença isolada do tipo *A criança gosta de doces e guloseimas*, é concebível que boa parte dos participantes tenha assumido (inconscientemente) um contexto discursivo prévio no qual se discutiam as diferenças entre o paladar das crianças e o dos adultos. Em outras palavras, é provável que, em muitos casos, a ausência da sequência contrastiva não tenha sido suficiente para suprimir a evocação do cenário conceptual associado ao SN genérico definido – uma hipótese que ajuda a explicar, inclusive, a alta incidência de sentenças plenamente aceitáveis (grau 5) na condição sem contraste.

É razoável especular, portanto, que o efeito seja mais forte do que o que foi possível observar neste experimento – e que a detecção de um efeito apenas moderado tenha sido artifício do desenho experimental. Para confirmar essa especulação, são bem-vindos experimentos que permitam flagrar de forma mais direta as representações conceptuais associadas aos diferentes padrões sintáticos – por exemplo, uma tarefa de escolha forçada com imagens. Para os nossos propósitos, a vantagem desse tipo de desenho é sua menor vulnerabilidade ao fenômeno da acomodação pragmática, que pode interferir na interpretação de sequências linguísticas. Seja como for, ainda que alguma forma de acomodação pragmática tenha de fato sido realizada por muitos sujeitos, e com um elevado número de estímulos, está claro que ela não foi suficiente para *anular* a diferença

¹⁷Essa tendência dá origem ao conhecido fenômeno de *acomodação de pressuposição* (LAMBRECHT, 1994): se alguém entra em um ambiente e ouve o enunciado “Foi o Diego que fez o gol do Flamengo”, provavelmente não poderá evitar a suposição de que a conversa anterior girava em torno da definição da autoria do gol do Flamengo.

na distribuição dos graus de aceitabilidade dos estímulos entre as duas condições experimentais. A rigor, o fato de o efeito do contraste ter sido observado mesmo numa situação propícia à acomodação pragmática pesa fortemente em favor da hipótese defendida aqui.

Um segundo ponto relevante diz respeito à existência de possíveis explicações alternativas para as diferenças de aceitabilidade encontradas. É possível notar que os estímulos das duas condições se distinguem não apenas quanto à contrastividade mas também quanto à informatividade – especificamente, os estímulos da condição com contraste veiculam uma maior quantidade de informações¹⁸. Para avaliar se os dados revelam a existência de um efeito de informatividade independente do efeito de contraste, um segundo experimento está atualmente em desenvolvimento.

5 Considerações finais

Este trabalho apresentou resultados experimentais compatíveis com uma análise cognitivista acerca da semântica de SNs singulares genéricos. A partir da caracterização do sintagma nominal estabelecida por Langacker (1991), postulamos a existência de uma diferença conceptual entre o SN genérico definido e o SN genérico indefinido (com determinante zero) do PB. Com base em uma tarefa de julgamento de aceitabilidade, foi possível confirmar a previsão de que a presença de um segundo SN genérico, com função contrastiva, na mesma sequência textual eleva significativamente a aceitabilidade de enunciados com SN genérico definido. Esse resultado foi interpretado como evidência de que, no PB, SNs genéricos definidos evocam representações conceptuais distintas, conforme indicado nas Figuras 7 a 10.

O tratamento experimental de constructos teóricos associados à Gramática Cognitiva é um campo ainda pouco explorado, principalmente no que se refere ao PB. Nesse sentido, a presente pesquisa abre um caminho promissor de investigação, que pode respaldar empiricamente a associação entre padrões sintáticos e estruturas conceptuais. Com relação à caracterização dos sintagmas nominais genéricos do PB, os resultados aqui encontrados motivam o aprofundamento da proposta a partir de novos desenhos experimentais, ao mesmo tempo em que sugerem a ampliação da análise para incluir outros tipos de SNs genéricos do PB.

¹⁸ Agradecemos a um dos pareceristas anônimos por essa observação.

Referências

BROCCIAS, C; HOLLMANN, W. B. Do we need summary and sequential scanning in (Cognitive) grammar? *Cognitive Linguistics*, De Gruyter, v. 18, n. 4, p. 487-522, 2007.

FAUCONNIER, G. *Mental spaces*. Cambridge: Cambridge University Press, 1994. <https://doi.org/10.1017/CBO9780511624582>

HAWKINS, J. *Definiteness and indefiniteness: a study in reference and grammaticality prediction*. London: Croom Helm., 1978.

HART, C. *Discourse, grammar and ideology: functional and cognitive perspectives*. London; New Delhi; New York; Sydney: Bloomsbury, 2014.

HOLLMANN, W. B. Nouns and verbs in Cognitive Grammar: where is the 'sound' evidence? *Cognitive Linguistics*, De Gruyter, v. 24, n. 2, p. 275-308, 2013.

JAMIESON, S. Likert scales: how to (ab)use them. *Medical Education*, Wiley Online Library, v. 38, n. 12, p. 1217-1218, 2004.

KRIFKA, M.; PELLETIER, F. J.; CARLSON, G.; TER MEULEN, A.; CHIERCHIA, G.; LINK, G. Genericity: an introduction. In: CARLSON, G.; PELLETIER, F. J. (Ed.). *The generic book*. Chicago; London: The University of Chicago Press, 1995. p. 1-124.

LAMBRECHT, K. *Information structure and sentence form: topic, focus and the mental representation of discourse referents*. Cambridge: University Press, 1994. <https://doi.org/10.1017/CBO9780511620607>

LANGACKER, R. W. *Foundations of cognitive grammar: theoretical prerequisites*. Stanford: Stanford University Press, 1987.

LANGACKER, R. W. *Foundations of cognitive grammar: descriptive application*. Stanford: Stanford University Press, 1991.

LANGACKER, R. W. *Investigations in Cognitive Grammar*. Berlin; New York: Mouton de Gruyter, 2009. <https://doi.org/10.1093/oxfordhb/9780199544004.013.0005>

LANGACKER, R. W. *Essentials of Cognitive Grammar*. Oxford: Oxford University Press, 2013.

MICHAELIS, L. A. Sign-Based Construction Grammar. In: HOFFMANN, T.; TROUSDALE, G. (Ed.). *The Oxford Handbook of Construction Grammar*. Oxford: University Press, 2013.

MÜLLER, A. The semantics of generic quantification in Brazilian Portuguese. *Probus*, Berlin, v. 14, n. 2, p. 279-298, 2002a. <https://doi.org/10.1515/prbs.2002.011>

MÜLLER, A. Genericity and the denotation of common nouns in Brazilian Portuguese. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, PUC-SP, v. 18, n. 2, p. 287-308, 2002b. <https://doi.org/10.1590/S0102-44502002000200005>

PEETERS, B. Does Cognitive Linguistics live up to its name? In: DIRVEN, R.; HAWKINS, B.; SANDIKCIOGLU, E. (Ed.) *Language and ideology: theoretical cognitive approaches*. Amsterdam/New York: John Benjamins, 2001. p. 83-106. <https://doi.org/10.1075/cilt.204.04pee>

REA, L. M.; PARKER, R. A. *Designing and conducting survey research*. San Francisco: Jossey-Bass, 1992.

SAVIC, M. *Politeness through the prism of requests, apologies and refusals: a case of advanced Serbian EFL learners*. Cambridge: Cambridge Scholars Publishing, 2014.

STEFANOWITSCH, A. Cognitive linguistics as a cognitive science. In: CALLIES, M.; KELLER, W. R.; LOHÖFER, A. (Ed.) *Bidirectionality in the cognitive sciences*. Amsterdam; Philadelphia: John Benjamins, 2011. p. 296-309. <https://doi.org/10.1075/hcp.32.16ste>

WESTERA, M.; BRASOVEANU, A. Ignorance in context: The interaction of modified numerals in QUDs. In: SEMANTICS AND LINGUISTIC THEORY CONFERENCE, 24., 2014, New York. *Proceedings...* New York: New York University, 2014. p. 414-431.

ANEXO 1

Estímulos críticos do Script A

1. A criança gosta de doces e guloseimas.
2. O futebol tem muitos adeptos no Brasil.
3. O carro só serve para atrapalhar a vida.
4. A batata frita engorda.
5. O chocolate dá espinha. Já a maçã faz bem para a pele.
6. O notebook ocupa muito espaço na mala. Já o tablet é leve e mais prático de carregar.
7. O casaco é totalmente desnecessário no inverno do Rio de Janeiro. Enquanto isso, a regata é um item comum em todas as estações.
8. O leite agrava a gastrite. Por outro lado, o chá de hortelã melhora problemas estomacais.

Estímulos críticos do Script B

1. O chocolate dá espinha.
2. O leite agrava a gastrite.
3. O notebook ocupa muito espaço na mala.
4. O casaco é totalmente desnecessário no inverno do Rio de Janeiro.
5. O futebol tem muitos adeptos no Brasil. Já o golfe não é muito popular.
6. A criança gosta de doces e guloseimas. Já o adulto tende a ter um paladar mais diversificado.
7. A batata frita engorda. Já a fruta é indicada para quem faz dieta.
8. O carro só serve para atrapalhar a vida. O metrô, por outro lado, facilita a locomoção das pessoas.

Estímulos distratores

1. Todo ônibus possui assentos preferenciais para idosos e gestantes.
2. Florestas são adequadas para se fazer trilhas.
3. Todo instrumentista possui instrução formal. Dessa maneira, há músicos que não possuem treinamento algum.
4. Uma boa parte dos avós gosta de contar histórias para os seus netos.
5. Todo quarto de criança é bagunçado.
6. As famílias costumam se encontrar aos domingos.
7. Os gatos são muito amigáveis. Entretanto, não têm medo das pessoas.

8. Os jogadores de futebol são muito famosos no Brasil. Mas os atletas de hóquei não costumam ter muitos fãs.
9. As sorveterias vendem mais em dias quentes. Assim, lucram mais no inverno.
10. Roqueiro costuma usar roupas pretas e pulseiras metálicas.
11. Os cantores líricos cantam apenas música sertaneja.
12. Lanchonete é um bom lugar para se fazer refeições rápidas.
13. As cidades grandes são muito barulhentas. Mesmo assim, são lugares recomendados para se morar
14. Flash é o herói mais rápido dos quadrinhos. Dessa forma, ele pode correr mais rápido que a velocidade da luz.
15. Cidades quentes são péssimas para se encontrar gorros e agasalhos.
16. Os escritores costumam não gostar de ler livros.

ANEXO 2

Você está prestes a participar de um experimento linguístico muito simples. Leia as instruções a seguir com bastante atenção e responda ao que for solicitado.

Imagine que você está ouvindo um pai transmitir ao seu filho uma série de ensinamentos sobre diversos assuntos e seres. Na sequência, você verá algumas das frases produzidas por ele. Sua tarefa consiste simplesmente em julgar se elas soam aceitáveis e naturais no português. Para isso, considere a seguinte escala:

1. Totalmente inaceitável
2. Pouco aceitável
3. Medianamente aceitável
4. Muito aceitável
5. Totalmente aceitável

Lembre-se: você não está sendo avaliado e não deve levar em consideração as regras gramaticais aprendidas na escola. Ao atribuir uma nota para cada frase (1, 2, 3, 4 ou 5), considere apenas a maneira como elas *soam* aos seus ouvidos. Em outras palavras, faça o seu julgamento utilizando apenas a sua intuição de falante do português.

Um estudo experimental do processamento de metáforas do português brasileiro

An experimental study on Brazilian Portuguese metaphor processing

Antonio João Carvalho Ribeiro

Grupo de Estudos e Laboratório em Psicolinguística Experimental – GEPEX
Universidade Federal Fluminense, Niterói, Rio de Janeiro / Brasil
ajcarvalhoribeiro@gmail.com

Adiel Queiroz Ricci

Grupo de Estudos e Laboratório em Psicolinguística Experimental – GEPEX
Universidade Severino Sombra, Vassouras, Rio de Janeiro / Brasil
adielricci@gmail.com

Resumo: Um estudo experimental do processamento psicolinguístico de metáforas nominais do português brasileiro (X é um Y), p. ex., “Irene é um furacão”, foi realizado com o objetivo de evidenciar, a partir de tempos de leitura (*RTs*), a compreensão de expressões familiares, *high-apt* (“bem construídas”) e cujo veículo se acha convencionalizado. Na primeira fase da pesquisa, realizaram-se dois *norming studies* (“estudos normativos”) com vistas ao ranqueamento de metáforas nominais (p. ex. “Algumas mulheres são furacões”) em relação a familiaridade, *aptness* (“adequação”) e convencionalidade. Na segunda fase da pesquisa, um experimento de leitura automonitorada (*self-paced, non-cumulative, moving-window reading*) foi conduzido, recorrendo, para a composição dos estímulos, às metáforas, p. ex., “Irene é um furacão”, que alcançaram, nos estudos normativos da primeira fase, *ratings* (ou “classificações”) de “muito familiares”, “*very high-apt*” e “altamente convencionalizadas”. Evidências do português brasileiro em favor do processamento direto de metáforas foram obtidas, conforme preconiza o modelo de *Class-*

inclusion, de Glucksberg e Keysar (1990), pois não se revelaram diferenças significativas entre os *RTs* médios nas três condições: “metáfora”, “literal” e “declaração literal de inclusão em classe”, em contraposição aos achados de Janus e Bever (1985), que observaram tempos de leitura de metáforas novas significativamente maiores do que os de expressões literais, conforme as predições do Modelo Pragmático Padrão de processamento indireto.

Palavras-chave: processamento psicolinguístico da metáfora; *class-inclusion*; *dual reference*; *norming studies*; compreensão de linguagem figurada; português brasileiro.

Abstract: An experimental study on the psycholinguistic processing of Brazilian Portuguese attributive metaphors (X is a Y), e.g., “*Irene é um furacão*” (“Irene is a hurricane”), was carried out with the aim of highlighting, from reading times (RTs), the understanding of familiar, high-apt (“well-built”) expressions and the vehicle of which is conventionalized. In the first phase of the research, two norming studies were carried out aiming at the ranking of attributive metaphors, e.g. “*Algumas mulheres são furacões*” (“Some women are hurricanes”), regarding familiarity, aptness and conventionality. In the second phase of the research, a self-paced, non-cumulative, moving-window reading experiment was conducted, using, for the composition of the stimuli, the metaphors, e.g., “*Irene é um furacão*”, that have reached, in the normative studies of the first phase, ratings of “very familiar”, “very high-apt” and “highly conventionalized”. Brazilian Portuguese evidences in favor of direct processing of metaphors were obtained, as recommended by the Class-inclusion model of Glucksberg and Keysar (1990), since there were no significant differences between the mean RTs in the three conditions: “Metaphor”, “Literal” and “Literal Declaration of Class Inclusion”. In contrast to the findings of Janus and Bever (1985), who observed reading times of new metaphors significantly larger than those of literal expressions, according to the predictions of the Standard Pragmatic Model of indirect processing.

Keywords: psycholinguistic processing of metaphor; class-inclusion; dual reference; norming studies; figurative language comprehension; brazilian portuguese.

Recebido em: 8 de dezembro de 2017.

Aprovado em: 16 de março de 2017.

1 Introdução

Este artigo pretende inserir o português brasileiro (PB) na literatura sobre compreensão de metáforas, polarizada, segundo Glucksberg (2003), entre modelos de processamento indireto, cujos pressupostos se têm revelado hipóteses testáveis, e processamento direto, com base em evidências empíricas da interferência das variáveis “Familiaridade”, “*Aptness*” (“eficácia”, “adequação”) e “Convencionalidade”. Para tanto, reportamos aqui os resultados da pesquisa em que nos envolvemos nos últimos dois anos, investigando a compreensão de metáforas que se acham convencionalizadas no PB, sugestivos de processamento direto segundo o modelo de *Class-inclusion* de Glucksberg e Keysar (1990).¹

Na fase inicial da pesquisa, foram realizados dois pré-testes – *norming studies* nos quais os participantes ranquearam expressões metafóricas, por exemplo, “Algumas mulheres são furacões”, em Familiaridade, *Aptness* e Convencionalidade.² Na literatura da área, são chamados de *norming studies* (BLASKO; CONNINE, 1993; BOWDLE; GENTNER, 2005; DULCINATI *et al.*, 2014), ou *norming pre-tests* (JONES; ESTES, 2006), ou ainda *rating experiments* (THIBODEAU; DURGIN, 2011), ou, simplesmente, de *rating tasks* (HARRIS, 1976), os instrumentos concebidos para avaliar expressões metafóricas. O objetivo dos ranqueamentos foi o de selecionar expressões, entre as que receberam os *ratings* mais altos em ambos, como norteadoras da construção dos estímulos do experimento de leitura automonitorada conduzido na fase subsequente. Por exemplo, os *ratings* altos atribuídos a “Algumas mulheres são furacões” basearam a construção do estímulo “Irene é um furacão” que, inserido em contextos (parágrafos) apropriados, permitiria interpretação metafórica (se Irene é uma mulher agitada) ou literal (se Irene é o furacão que atingiu o Caribe e os EUA em 2011).

A opção por utilizar metáforas nominais – do tipo X é um Y – em que X é o tópico e Y é o veículo, nos exemplos acima, respectivamente,

¹ A pesquisa em questão, orientada pelo primeiro autor, foi conduzida pelo segundo, como parte dos requisitos de obtenção, por esse último, do título de Mestre em Linguística (RICCI, 2016). Aprovação pelo CEP da USS – RJ, CAAE 50572215.80000.5290, parecer 1.412.627, em 22 fev. 2016.

² Os resultados dos referidos *norming studies* foram apresentados no XXXI ENANPOLL (RIBEIRO *et al.*, 2016)

“Algumas mulheres” e “Irene” / “furacões” e “furacão”, deve-se ao fato de a maioria dos experimentos que se acham descritos na literatura focalizar essa forma (DULCINATI *et al.*, 2014, p. 72).

Finalmente, a decisão de propor uma tarefa de leitura automonitorada de expressões do PB em contextos diversos visou a superar as fragilidades metodológicas apontadas por Janus e Bever (1985) nos experimentos conduzidos até então – sustentando (cf. CLARK; LUCY, 1975; GIBBS JR., 1979, 1981) ou refutando (cf. ORTONY *et al.*, 1978; GLUCKSBERG; GILDEA; BOOKIN, 1982) o que se convencionou chamar de Modelo Pragmático Padrão de processamento indireto da metáfora (cf. SEARLE, 1993 [1979]; GRICE, 1975) – nos quais não se preveem contextos, ou tempos de leitura são medidos ao final da apresentação dos estímulos. Janus e Bever (1985) mediram tempos de leitura dos fragmentos críticos de expressões contextualizadas, isto é, dos sintagmas nominais que contêm o veículo – *locus* do “Problema do Reconhecimento” do significado metafórico – e observaram que metáforas “novas” tomaram, em média, tempos de leitura significativamente maiores do que expressões equivalentes de significado literal. Sendo assim, ofereceu-se, na literatura, o espaço que este estudo no PB busca ocupar: adotando as linhas mestras do experimento que foi conduzido por Janus e Bever (1985), oferecer achados cronométricos em proveito da tese do processamento direto de metáforas familiares, *high-apt* e cujo veículo acha-se convencionalizado.

2 Revisão da literatura

Tradicionalmente, na esteira da visão aristotélica, de que a metáfora é desviante da linguagem ordinária (GARRET, 2007), a interpretação metafórica é considerada como opcional (GLUCKSBERG, 2003; BOWDLE; GENTNER, 2005). No que se convencionou chamar de Modelo Pragmático Padrão (SPM), o processamento da metáfora (bem como o da linguagem figurada em geral) se dá em três estágios. Inicialmente, o ouvinte / leitor deriva o significado literal da expressão. No segundo estágio, ele verifica se a interpretação é apropriada no contexto do enunciado e deve ser aceita, ou não, como significado pretendido pelo locutor. Se implausível, a representação inicial é rejeitada, e chega-se ao terceiro estágio, o da busca por uma interpretação metafórica. Portanto, a malograda interpretação inicial do significado literal funciona como

o “gatilho” que dispara a busca por um significado metafórico que dê conta do proferimento (SEARLE, 1993 [1979], p. 89-90). Sendo assim, metáforas nominais como, por exemplo, *Some jobs are jails* (“Alguns empregos são prisões”) (exemplo retirado de GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998, 2003) seriam reconhecidas como asserções que ferem a máxima da qualidade (GRICE, 1975), e uma maneira de compreendê-las passaria por sua conversão em símile, *Some jobs are like jails* (“Alguns empregos são como prisões”), declarações literais verdadeiras, posto que duas coisas quaisquer sempre podem ser semelhantes de inúmeras maneiras.

Entretanto, Glucksberg, Gildea e Bookin (1982) questionam a prioridade do significado literal/ opcionalidade da interpretação metafórica, pressuposto do modelo SPM, a partir das evidências que obtiveram em uma série de experimentos de verificação de sentenças inspirados no teste de Stroop (1935). Stroop (1935) evidenciou a competição entre processos atencionais que permite capturar o automatismo das fases iniciais do processamento da leitura, traduzido em erros ou maiores tempos de reação ao se responder em que cor uma palavra (por exemplo, vermelho) está escrita, quando ela se acha impressa em outra cor (por exemplo, verde). Glucksberg, Gildea e Bookin (1982.) alegam que as metáforas que eles apresentaram aos participantes, p. ex., *some surgeons are butchers* (“alguns cirurgiões são açougueiros”) tomaram mais tempo para serem julgadas como falsas do que as metáforas “embaralhadas”, p. ex., *some jobs are butchers* (“alguns empregos são açougueiros”), porque os significados metafóricos disponíveis nas metáforas, mas não nas metáforas embaralhadas, interferiram no julgamento do significado literal dessas expressões. Recorrendo à lógica do efeito Stroop, ou da “compulsão” por ler a palavra quando se é instado a dizer o nome da cor em que a ela está escrita, Glucksberg, Gildea e Bookin (1982) refutam o modelo SPM, segundo o qual se cogitam interpretações metafóricas apenas quando os significados literais se mostram intratáveis, sustentando que interpretações metafóricas não podem ser inibidas ou, em suas próprias palavras, que o significado metafórico não pode ser ignorado.

De acordo com Glucksberg e colaboradores (GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998; GLUCKSBERG, 2003), falham as diversas teorias (cf. TVERSKY, 1977; ORTONY, 1993 [1979]) que tratam metáforas como comparações implícitas porque, em comparações literais, a equação “X é como um Y” é bidirecional, p. ex.: *copper is like tin*

ou *tin is like cooper* (“cobre é como estanho” ou “estanho é como cobre”), *coffee is like tea* ou *tea is like coffee* (“café é como chá” ou “chá é como café”), enquanto, em metáforas, os elementos não podem ser comutados, p. ex.: *some jobs are jails*, mas não **some jails are jobs* (“alguns empregos são prisões”, mas não “*algumas prisões são empregos”), *sermons are sleeping pills*, mas não **sleeping pills are sermons* (“sermões são soníferos”, mas não “*soníferos são sermões”). A aparente reversibilidade dos elementos da metáfora produz uma nova expressão, p. ex.: *some surgeons are butchers* (“alguns cirurgiões são açougueiros”) e *some butchers are surgeons* (“alguns açougueiros são cirurgiões”) diferem nas bases que sustentam a construção de cada uma delas.

De acordo com Glucksberg e colaboradores (GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998), metáforas são o que aparentam ser: declarações de inclusão em classe, tanto quanto asserções literais de inclusão em categorias. A metáfora *sermons are sleeping pills* (“sermões são soníferos”), por exemplo, comporta-se como uma declaração de inclusão de *sermons* (“sermões”) na categoria metafórica *sleeping pills* (“soníferos”), tal qual a asserção *a tree is a plant* (“uma árvore é um vegetal”) inclui *tree* (“árvore”) na classe *plant* (“vegetal”) (GLUCKSBERG; KEYSAR, 1990, p. 12). É evidência, segundo eles, de que, em metáforas nominais, “X é um Y” expressa relação de inclusão em classe, e não de igualdade, a possibilidade de parafrasear, sem problemas, comparações metafóricas em metáforas (na forma paradigmática), p. ex.: *my job is like a jail* → *my job is a jail* (“meu emprego é como uma prisão → meu emprego é uma prisão”); ao passo que é impossível parafrasear comparações literais em declarações verdadeiras, p. ex.: *bees are like hornets* → **bees are hornets* (“abelhas são como vespas → *abelhas são vespas”) (GLUCKSBERG; KEYSAR, 1990, p. 7-11) ou *my lawyer was like a shark* → *my lawyer was a shark* (“meu advogado é como um tubarão → meu advogado é um tubarão”), mas não *barracudas are like sharks* → **barracudas are sharks* (“barracudas são como tubarões → *barracudas são tubarões”) (GLUCKSBERG, 1998, p. 41).

De acordo com o modelo teórico de Glucksberg e Keysar (1990) – *Class-inclusion* –, metáforas são compreendidas como declarações de inclusão em classes, ou categorias, nas quais os veículos metafóricos têm o que ele chama de *Dual Reference* (“dupla referência”). Por exemplo, *shark* refere um predador marinho (ou uma categoria deles) e também uma categoria abstrata de entes predatórios em geral à qual empresta o

nome. Em *my lawyer was a shark* (“meu advogado era um tubarão”) o veículo da expressão, (*shark* / “tubarão”) estabelece uma relação taxonômica entre predador marinho e advogado, ambos abrigados em categoria abstrata de entes predatórios em geral que recebe o nome do membro prototípico: *shark* / “tubarão”. Em síntese, metáforas cujos veículos referem categorias metafóricas convencionalizadas são processadas diretamente, tanto quanto asserções literais de inclusão em categoria, nos termos do modelo de *Class-inclusion / Dual reference*, em que a “dupla referência” provê explicação tanto para a “não reversibilidade da metáfora” quanto para a possibilidade de parafrasear comparações metafóricas em declarações, o que permite distinguir metáforas de comparações (GLUCKSBERG, 2003, p. 95). A par desses fatos, Glucksberg (1998, 2003) rechaça a hipótese de que metáforas seriam processadas a partir da conversão em símiles, amparado nos achados empíricos de Johnson (1996), de que metáforas na forma de declaração, p. ex.: *my lawyer is a shark* (“meu advogado é um tubarão”) foram compreendidas mais rapidamente que símiles, *my lawyer is like a shark* (“meu advogado é como um tubarão”), embora essas contenham uma palavra (*like* / “como”) mais do que aquelas.

Ao legitimar a possibilidade de o membro prototípico designar a categoria metafórica, Glucksberg e Keysar (1990, p. 8) alegam que diversas línguas se valem, rotineiramente, da estratégia para rotular categorias supraordenadas. Na língua de sinais americana, utilizam-se sinais para itens básicos do mobiliário, *chair* (“cadeira”), *table* (“mesa”), *bed* (“cama”) para referir categoria *furniture* (“móvel”). No birmanês, falado no sudeste da Ásia (Birmânia, Tailândia, Malásia e Singapura) recorre-se à dupla função de nomes que referem objetos altamente prototípicos e classes (DENNY, 1986; CRAIG, 1986 *apud* GLUCKSBERG; KEYSAR, 1990). Até mesmo em línguas não classificadoras, como o hebraico e o inglês, identificam-se exemplos da utilização de nomes prototípicos para designar categorias, como o sobrenome de um acusado de crimes de guerra, *Demjanjuk*, que, em Israel e artigos de jornais norte-americanos tornou-se a palavra para identificar “indivíduo comum capaz de cometer atos inomináveis” (SHINOFF, 1987 *apud* GLUCKSBERG; KEYSAR, 1990) –, ainda que John Demjanjuk, não tenha sido julgado, definitivamente, culpado dos crimes atribuídos ao sádico guarda do campo de concentração de Treblinka, na Polônia, chamado de “Ivan o Terrível”. Em grande número

de línguas indígenas do sudoeste dos Estados Unidos, há fartos exemplos do uso ocasional de nomes de membros prototípicos para rotular as categorias em que se incluem. Na língua hopi, o nome da espécie mais abundante, *cottonwood*, pode designar todas as árvores chamadas de *deciduous trees*, ou caducifólias ou de folhas “caducas”, que perdem completamente as folhas no outono e no inverno (TRAGER, 1938 *apud* GLUCKSBERG; KEYSAR, 1990). E em shoshoni, a palavra para “águia” também é usada para aves de grande porte em geral (HAGE; MILLER, 1976 *apud* GLUCKSBERG; KEYSAR, 1990). Segundo Glucksberg e Keysar (1990), o princípio subjacente aos exemplos de dupla referência acima – tomar o membro que é prototípico para referir a categoria – explica o que se dá nas metáforas: toma-se o veículo para referir a categoria metafórica, abstrata, além de referir o ente real. Há casos em que a dupla referência do veículo das metáforas fica explícita, como, por exemplo, na fala de um espectador presente ao julgamento de John Demjanjuk, transcrita por Glucksberg e Keysar (1990, p. 8): *I know his name is Demjanjuk, but I don't know if he is a Demjanjuk* (“Eu sei que o nome é Demjanjuk, mas não sei se ele é um Demjanjuk”). E no exemplo oferecido por Glucksberg (2003, p. 94): *Cambodia was Vietnam's Vietnam* (“O Camboja foi o Vietnã do Vietnã”), em que, em inglês, a primeira menção a Vietnã refere o país, e a segunda menção (de Vietnã), a categoria abstrata supraordenada das intervenções militares desastrosas decorrentes do conflito armado no sudeste da Ásia (e vice-versa em português).

Finalmente, no mapeamento do campo de pesquisas em psicolinguística da metáfora até então explorado, Glucksberg (2003) enfatiza as evidências dos experimentos de Blasko e Connine (1993) de *crossmodal priming* (no caso, *priming* semântico associado a uma tarefa de decisão lexical), de que *ratings* altos de familiaridade e *aptness* aceleram a ativação do significado metafórico do veículo de metáforas. Com isso, Glucksberg (2003) arrola, além da convencionalidade do veículo, o papel das variáveis Familiaridade e *Aptness* (“adequação ou eficácia da expressão”) no acesso rápido ao significado metafórico e, portanto, tributárias do processamento direto de expressões metafóricas.

3 Experimento 1: *norming study* de Familiaridade e Convencionalidade

O Experimento 1 objetivou avaliar a familiaridade dos participantes em relação às metáforas que foram construídas, para tal fim, pelo pesquisador, sendo familiaridade aqui entendida como familiaridade subjetiva, segundo a definição de Blasko e Connine (1993, p. 305): *the perceived experience with the metaphor* (“a sensação da experiência com a metáfora”)³ Para além da obtenção de ratings de familiaridade das metáforas – em familiar e não familiar – estimou-se a convencionalidade das expressões, cujos veículos tiveram o significado figurado colhido em dicionários de prestígio do PB.⁴ Para tanto, cada um dos participantes escolheu, entre três palavras, aquela que, de acordo com ele(a), capturava o significado que a expressão pretendia comunicar. A partir das respostas, verificou-se se o veículo (Y) evocara o significado figurado ou por extensão registrado nos dicionários consultados, permitindo, ou não, inferir se o mesmo se achava, ou não, convencionalizado no repertório dos indivíduos da amostra.

Participantes: 81 estudantes da Universidade Severino Sombra, em Vassouras – RJ, com idade entre 18 e 56 anos, sendo 16 do sexo masculino e 65 do feminino, todos falantes nativos de PB, jamais tendo colaborado em experimentos sobre compreensão de metáforas, participaram voluntariamente.

Materiais: 84 metáforas nominais (X é um Y) foram construídas, pelo pesquisador, sem repetições de veículos (Y) ou tópicos (X). Os tópicos das expressões foram modificados por “alguns(mas)” e tanto os tópicos (X) quanto os veículos (Y) são substantivos concretos, de mesmo gênero, por exemplo: “Alguns carros são abacaxis”. Com exceção de alguns tópicos (X), que são comuns de dois gêneros e alguns veículos, sobrecomuns, p. ex.: “Alguns motoristas são lesmas”. Os significados figurados de todos os veículos (Y) acham-se registrados em pelo menos

³ Em oposição a familiaridade objetiva, estimada com base na frequência metafórica, utilizando a ferramenta de pesquisa Google com corpus (THIBODEAU; DURGIN, 2011, p. 209).

⁴ AULETE. *Aulete Digital*: o dicionário da língua portuguesa na internet; HOUAISS; VILLAR. *Dicionário Houaiss da Língua Portuguesa*, 3. ed. (2009); FERREIRA. *Novo Aurélio século XXI*: o Dicionário da Língua Portuguesa (1999); e a 6ª edição do FERREIRA. *Miniaurélio: o Minidicionário da Língua Portuguesa* (2004).

dois dos dicionários consultados. E, tomadas literalmente, todas as expressões construídas são falsas.

Procedimentos: Em dias pré-determinados, grupos de participantes de, no máximo, 19, foram acomodados em uma das salas do Núcleo de Ensino Semipresencial da USS – RJ⁵ e, após firmarem o Termo de Consentimento Livre e Esclarecido, realizaram a tarefa proposta, em média, em 15 minutos. Cada um dos participantes sentou-se em frente a uma tela LCD de 17” que mostrava o formulário Google (Google Chrome sobre Windows 7) que continha as perguntas que deveriam responder (cada uma das telas conectada a um PC com teclado com fio ABNT2 e mouse USB óptico, com fio, de três botões e *scroll*). No cabeçalho de cada um dos blocos do formulário, imediatamente acima da expressão que tinham que julgar (familiar ou não familiar) os participantes podiam ler as instruções acerca do procedimento a ser seguido: “As expressões abaixo lhe são familiares?” “Você já as leu ou ouviu?” E, continuando: “Marque SIM (se for familiar) ou NÃO (se não for familiar)”. E “Depois, tendo, ou não, ouvido ou lido a expressão, escolha uma palavra que você acredita que capture o seu significado”.

Assim, foi possível estimar a convencionalidade do veículo (Y) de cada uma das metáforas, cujo significado figurado achava-se entre as opções, além de avaliar a qualidade da participação dos componentes da amostra, pois as outras opções referiam o significado literal da expressão ou o oposto do metafórico. Na distribuição das múltiplas escolhas, cuidou-se de alternar, dentro do formulário, as posições em que apareciam as opções (cf. excerto do formulário, na Figura 1). Dar respostas a todas as perguntas era obrigatório, impossibilitando que fossem deixadas em branco quando do envio do formulário via acesso à internet. E, ao final de cada sessão, os participantes puderam conhecer os objetivos do experimento e ter esclarecidas dúvidas que eventualmente tivessem.

⁵ Núcleo de Ensino Semipresencial da Universidade Severino Sombra, em Vassouras – RJ.

Figura 1 – Excerto do formulário de Familiaridade

FAMILIARIDADE
 As expressões abaixo lhe são familiares? Você já as leu ou ouviu?
 Marque SIM (se for familiar) ou NÃO (se não for familiar).
 Depois, tendo, ou não, ouvido ou lido a expressão, escolha uma palavra que você acredita que capture o seu significado.

1. Alguns carros são abacaxis. *
 (Esta expressão lhe é familiar?)

SIM
 NÃO

Agora escolha uma palavra que capture o significado da expressão acima: *

COMPLICAÇÃO
 FACILIDADE
 FRUTO

3.1 Experimento 2: *norming study* de *Aptness*⁶ e Convencionalidade

O Experimento 2 objetivou avaliar em *aptness* as expressões do conjunto de 84 metáforas cuja familiaridade foi julgada no Experimento 1. Atendendo à definição de *aptness* mais amplamente aceita: *the extent to which the statement captures important features of the topic* (cf. CHIAPPE *et al.*, 2003, p. 97) (“o quanto a declaração captura as características importantes do tópico”),⁷ foi pedido aos participantes que dessem notas de 1 a 5 a cada uma das expressões, avaliando a sua eficácia em transmitir o significado que pretendem transmitir. Além disso, estimou-se, como foi feito no Experimento 1, a convencionalidade das expressões.

Participantes: 79 estudantes da Universidade Severino Sombra, em Vassouras – RJ, com idade entre 19 e 44 anos, sendo 28 do sexo masculino e 51 do feminino, todos falantes nativos de PB, jamais

⁶ A versão, em português, de *aptness*, aqui, bem como nos formulários apresentados aos participantes, “eficácia”, poderia ser “capacidade” ou “competência”, “jeito”, ou, até mesmo, “adequação” ou “a qualidade do que é apropriado ou adequado”.

⁷ Dulcinati *et al.* (2014, p. 74), advertem que se trata de uma propriedade da metáfora que tem recebido muitas definições diferentes.

tendo colaborado em experimentos sobre compreensão de metáforas, participaram voluntariamente.

Materiais: As mesmas 84 metáforas nominais (X é um Y) mostradas aos participantes do Experimento 1.

Procedimentos: Em dias pré-determinados, grupos de participantes de, no máximo, 19, foram acomodados em uma das salas do Núcleo de Ensino Semipresencial da USS – RJ⁸ e, após firmarem o Termo de Consentimento Livre e Esclarecido, realizaram a tarefa proposta, em média, em 15 minutos. Cada um dos participantes sentou-se em frente a uma tela LCD de 17” que mostrava o formulário Google (Google Chrome sobre Windows 7) que continha as perguntas que deveriam responder (cada uma das telas conectada a um PC com teclado com fio ABNT2 e mouse USB óptico, com fio, de três botões e *scroll*). No cabeçalho de cada um dos blocos do formulário, imediatamente acima da expressão que tinham que julgar, os participantes podiam ler as instruções acerca do procedimento a ser seguido: “Dê uma nota de 1 a 5 a cada uma das expressões abaixo, avaliando a sua eficácia em transmitir o significado que pretendem transmitir”, e “Depois, escolha uma palavra que você acredita que capture o seu significado”. Tal qual no Experimento 1, foi possível estimar a convencionalidade do veículo (Y) de cada uma das metáforas, cujo significado figurado achava-se entre as opções, além de avaliar a qualidade da participação dos componentes da amostra, pois as outras opções referiam o significado literal da expressão ou o oposto do metafórico. Na distribuição das múltiplas escolhas, cuidou-se de alternar, dentro do formulário, as posições em que apareciam as opções (cf. excerto do formulário, na Figura 2). Dar respostas a todas as perguntas era obrigatório, impossibilitando que fossem deixadas em branco quando do envio do formulário via acesso à internet. E, como no Experimento 1, ao final de cada sessão, os participantes puderam conhecer os objetivos do experimento e ter esclarecidas dúvidas que eventualmente tivessem.

⁸ Nas mesmas dependências em que foi realizado o Experimento 1.

Figura 2 – Excerto do formulário de *Aptness*

APTNESS
 Dê uma nota, de 1 a 5, a cada uma das expressões abaixo, avaliando a sua eficácia em transmitir o significado que pretendem comunicar.

Depois escolha uma palavra que você acredita que capture o significado da expressão.

1. Alguns carros são abacaxis. *

1 2 3 4 5

● ● ● ● ●

Agora escolha uma palavra que capture o significado da expressão acima: *

● COMPLICAÇÃO
 ● FACILIDADE
 ● FRUTO

3.2 Resultados dos *norming studies*

Os resultados dos 2 *norming studies* contabilizaram um total de 26.880 julgamentos das 84 expressões metafóricas que foram construídas para esse fim: 13.608 julgamentos feitos pelos 81 participantes do Experimento 1 a respeito da “Familiaridade” e da “Convencionalidade” das metáforas; e outros 13.272, feitos pelo grupo de 79 participantes do Experimento 2, que as ranquearam em “*Aptness*” e “Convencionalidade” – em ambos os casos, com a finalidade de compor o conjunto de estímulos do experimento de leitura automonitorada que foi realizado na fase seguinte da pesquisa. A seleção das metáforas de *rating* mais alto em “Convencionalidade” baseou-se na média das estimativas obtidas nos 2 experimentos.

Considerando-se o indivíduo, familiaridade deve ser vista como uma variável discreta, envolvendo exclusivamente dois fatores, ou “sentimentos”: um, de familiaridade com “algo”, e outro de não familiaridade com “esse algo”, se o que é percebido pelo indivíduo lhe parece, respectivamente, familiar ou não familiar. Porém, tomando grupos de indivíduos, é lícito admitir graus de familiaridade, tais como os *ratings* de “Familiaridade” obtidos entre os participantes do Experimento 1, que resultaram em metáforas “não familiares”, “pouco familiares”, “medianamente familiares”, “familiares” e “muito familiares” reunidos na Tabela 1.

Tabela 1 – *Ratings* de “Familiaridade” no Experimento 1

Ratings das expressões	Respostas SIM (%)⁹
Metáforas muito familiares	entre 80 e 100
Metáforas familiares	igual ou maior que 60 e menor que 80
Metáforas medianamente familiares	igual ou maior que 40 e menor que 60
Metáforas pouco familiares	igual ou maior que 20 e menor que 40
Metáforas não familiares	menor que 20

4 expressões, isto é, menos de 5% do total, foram julgadas familiares por menos da metade dos participantes: “Alguns vizinhos são sapos” (familiar a 41,98%), “Alguns carros são abacaxis” (40,74%), “Algumas explicações são véus” (39,51%) e a menos familiar ao grupo, “Alguns indivíduos são avestruzes” (37,04%). A grande maioria, 75 das 84 expressões (89% do total) foi julgada familiar por 60% ou mais dos participantes: desde aquelas que pareceram familiares a pouco mais da metade do grupo, por exemplo, “Alguns pensadores são águias” (60,49%) e “Algumas amigas são camaleões” (64,02%) às que soaram familiares a todos, p. ex.: “Algumas sogras são cobras” (100%) e “Algumas mulheres são furacões” (100%). Essa maioria abarcou as metáforas ranqueadas “familiares” e “muito familiares”, pré-requisito da seleção dos estímulos do experimento de leitura automonitorada a se realizar na fase seguinte da pesquisa.

Contudo, decidimos elevar o nível de corte para 85% de respostas positivas (SIM) nos julgamentos da familiaridade das expressões, preservando 59 (ou 70%) das 84 candidatas à seleção para compor o conjunto de estímulos do experimento de leitura automonitorada previsto. Sendo assim, as expressões metafóricas selecionadas foram julgadas familiares por 85% dos participantes, por exemplo, desde “Alguns políticos são raposas” (85,19%), passando por “Alguns professores são carrascos” (98,77%), até aquelas que alcançaram a unanimidade no grupo, p. ex.: “Alguns motoristas são lesmas” (100%) e “Alguns músicos são feras” (100%).

⁹ “As expressões abaixo lhe são familiares? Você já as leu ou ouviu? Marque SIM (se for familiar) ou NÃO (se não for familiar)” (cf. excerto do formulário de “Familiaridade”, Figura 1).

Tabela 2 – Ratings de Aptness no Experimento 2

<i>Ratings das expressões</i>	<i>Notas (de 1 a 5)¹⁰</i>
Metáforas <i>very high-apt</i>	5
Metáforas <i>high-apt</i>	4
Metáforas <i>moderate-apt</i>	3
Metáforas <i>low-apt</i>	2
Metáforas <i>very low-apt</i>	1

O ranqueamento em *Aptness*, de acordo com as notas de 1 a 5 dadas pelos participantes do Experimento 2 à eficácia das expressões, resultou nos *ratings*: *very low-apt*, *low-apt*, *moderate-apt*, *high-apt* e *very high-apt*, reunidos na Tabela 2.

Nenhuma das 84 expressões mereceu, dos participantes do Experimento 2, nota igual ou menor que 2 em *Aptness*, redundando em não se ter, no conjunto, metáforas de *ratings low-apt* e *very-low apt*. E apenas 6, ou 7% do total, receberam notas menores que 4, caracterizadas como *moderate-apt*, a saber: “Alguns carros são abacaxis” (com a média 3,94), “Alguns estudantes são traças” (com 3,89), “Algumas explicações são véus” (3,86), “Alguns vizinhos são sapos” (3,77) e as duas de *rating* mais baixo, “Alguns humanos são canários (3,65) e “Alguns indivíduos são avestruzes” (3,57). Em matéria de “eficácia”, a quase totalidade das expressões, 78 ou cerca de 93%, recebeu notas entre 4 e 5, circunscrevendo a maioria aos *ratings* mais altos, *high-apt* ou *very high-apt*, p. ex.: “Alguns viciados são zumbis” (com a média 4,08), “Alguns advogados são tubarões” (4,32), “Algumas estradas são serpentes” (4,35), “Alguns cunhados são malas” (4,72), “Algumas mães são leas” (4,82) e aquela que chegou mais perto de alcançar a nota máxima, “Algumas meninas são gatas” (4,91).

Apesar das notas altas dadas à maioria das expressões, aplicamos, aqui, o rigor adotado antes, elevando para 4,5 o nível de corte nos *ratings* de *Aptness* das metáforas, preservando então 61 das 84 (ou 73%) das candidatas a estímulos do experimento de leitura automonitorada previsto.

¹⁰ “Dê uma nota de 1 a 5 a cada uma das expressões abaixo, avaliando a sua eficácia em transmitir o significado que pretendem transmitir” (cf. excerto do formulário de “*Aptness*”, Figura 2).

Portanto, as expressões metafóricas selecionadas foram aquelas de *rating* igual ou maior do que 4,5, p. ex.: “Alguns policiais são gorilas” (4,54), “Alguns pais são bananas” (4,67), passando por “Alguns trabalhos são prisões” (4,71), “Algumas modelos são aviões” (4,8) até as que estiveram a ponto de alcançaram a unanimidade no grupo, tais como “Alguns lutadores são touros” (4,89) e “Alguns mestres são espelhos” (4,89).

As estimativas médias de “Convencionalidade”, obtidas a partir dos resultados dos Experimentos 1 e 2 traduziu-se nos *ratings*: “metáforas novas”, “metáforas pouco convencionalizadas”, “metáforas medianamente convencionalizadas”, “convencionalizadas” e “metáforas altamente convencionalizadas”, reunidos na Tabela 3. As estimativas com respeito à “Convencionalidade” coletadas nos Experimentos 1 e 2 foram reunidas em estimativas médias, ponderadas em relação ao número de participantes de cada um deles, com o objetivo de tirar proveito da totalidade das observações. A adoção do mesmo nível de corte (85%) utilizado para selecionar expressões com base em familiaridade, restringiu aqui os estímulos do experimento de leitura automonitorada previsto às expressões altamente convencionalizadas.

Tabela 3 – *Ratings* de “Convencionalidade” nos Experimentos 1 e 2

<i>Ratings</i> das expressões	Respostas congruentes (%)¹¹
Metáforas altamente convencionalizadas	entre 80 e 100
Metáforas convencionalizadas	igual ou maior que 60 e menor que 80
Metáforas medianamente convencionalizadas	igual ou maior que 40 e menor que 60
Metáforas pouco convencionalizadas	igual ou maior que 20 e menor que 40
Metáforas novas	menor que 20

Desse modo, foram selecionadas 43 expressões do conjunto original, com o objetivo de compor os estímulos do experimento de leitura automonitorada a se realizar na fase seguinte da pesquisa. Admitiram-se, nessa seleção: metáforas “muito familiares” (em ao menos 85% dos julgamentos dos participantes do Experimento 1), metáforas *very high-apt* (de *rating* igual ou superior a 4,5 obtido no ranqueamento do

¹¹ “[...] escolha uma palavra que você acredita que capture o significado da expressão” (cf. excertos dos formulários de “Familiaridade” e de “*Aptness*”, Figura 1 e Figura 2, respectivamente).

Experimento 2; e “altamente convencionalizadas” (em 85% ou mais das estimativas médias nas avaliações dos participantes dos Experimentos 1 e 2, todas reunidas na Tabela 4.

Tabela 4 – As 43 expressões selecionadas nos Experimentos 1 e 2

As 43 expressões selecionadas	Familiaridade	Aptness	Convencionalidade
Alguns políticos são raposas.	85,19%	4,67	97,45%
Algumas crianças são anjos.	98,77%	4,77	89,83%
Alguns chefes são toupeiras.	92,59%	4,66	94,92%
Alguns seguranças são armários.	93,83%	4,76	95,56%
Algumas modelos são aviões.	97,53%	4,8	98,73%
Alguns pais são bananas.	92,59%	4,67	97,45%
Alguns maridos são cavalos.	96,30%	4,68	100,00%
Algumas sogras são cobras.	100,00%	4,82	99,36%
Alguns meninos são diabos.	98,77%	4,67	96,84%
Algumas modas são febres.	96,30%	4,72	96,19%
Algumas senhoras são flores.	91,36%	4,68	99,37%
Algumas mulheres são furacões.	100,00%	4,81	98,11%
Algumas meninas são gatas.	100,00%	4,91	100,00%
Algumas adolescentes são girafas.	92,59%	4,63	99,36%
Alguns cordeiros são lobos.	91,36%	4,68	96,83%
Algumas mães são leoas.	97,53%	4,82	98,11%
Alguns motoristas são lesmas.	100,00%	4,85	98,73%
Alguns genros são sanguessugas.	91,36%	4,66	98,73%
Alguns avós são santos.	86,42%	4,65	87,97%
Algumas tias são vitrolas.	90,12%	4,72	98,09%
Algumas avós são corujas.	98,77%	4,73	98,11%
Alguns cunhados são malas.	97,53%	4,72	96,83%
Alguns músicos são feras.	100,00%	4,84	100,00%
Algumas motos são foguetes.	92,59%	4,85	99,37%
Alguns cozinheiros são porcos.	97,53%	4,85	94,33%
Alguns locutores são papagaios.	92,59%	4,72	96,20%
Algumas atitudes são máscaras.	85,19%	4,61	92,42%
Algumas provas são torturas.	98,77%	4,81	100,00%
Alguns botequins são chiqueiros.	95,06%	4,85	96,81%

Alguns professores são carrascos.	98,77%	4,87	96,20%
Algumas celebridades são antas.	88,89%	4,75	99,36%
Alguns fumantes são chaminés.	96,30%	4,86	97,47%
Alguns barracos são fornos.	93,83%	4,66	98,72%
Alguns caminhoneiros são tartarugas.	95,06%	4,85	98,72%
Algumas casas são fornalhas.	95,06%	4,72	97,47%
Algumas garotas são violões.	95,06%	4,8	97,44%
Alguns amigos são pilares.	87,65%	4,77	96,84%
Algumas torcidas são quadrilhas.	92,59%	4,63	89,89%
Algumas críticas são coices.	92,59%	4,72	94,30%
Alguns lutadores são touros.	93,83%	4,89	98,08%
Alguns mestres são espelhos.	93,83%	4,89	98,73%
Alguns trabalhadores são formigas.	92,59%	4,62	86,08%
Algumas vizinhas são moscas.	93,83%	4,51	95,53%

4 Experimento 3: leitura automonitorada de metáforas do PB

O experimento de leitura automonitorada (*self-paced, non-cumulative, moving-window reading*) objetivou evidenciar o processamento direto, sustentado pelo modelo teórico de *Class-inclusion* (GLUCKSBERG; KEYSAR, 1990), de metáforas do PB muito familiares, *very high-apt* e altamente convencionalizadas, selecionadas nos *norming studies* que realizamos, observando-se os tempos de leitura sobre o *locus* do “Problema do Reconhecimento” do significado metafórico, isto é, o veículo de cada uma das expressões (p. ex.: “um furacão” quando “Irene é um furacão” demandava interpretação metafórica). Foram registrados também os tempos de leitura de fragmentos correspondentes, p. ex.: “um furacão” quando “Irene é um furacão” tinha interpretação literal, e ainda os tempos de leitura de fragmentos-controle, p. ex.: “uma jovem” em declarações literais de inclusão em classe, no caso, “Irene é uma jovem”.

Participantes: 66 voluntários participaram do experimento: 20 homens e 46 mulheres, com idade média de 27 anos. 35 foram recrutados entre os estudantes do Curso de Graduação em Letras da Universidade Federal Fluminense e 3 eram funcionários administrativos da instituição, esses, com nível superior completo de escolaridade. 28 eram estudantes de diversos cursos da Universidade Severino Sombra (Enfermagem,

Engenharia de Produção, Engenharia Química, Medicina e Psicologia). Desses, 24 eram graduandos e 4 tinham Especialização.

Materiais: 9 trincas de parágrafos (transcritos no Apêndice), igualmente divididos em 11 fragmentos / sintagmas, foram preparadas para acomodar: no tipo 1: expressões do tipo “X é um Y” construídas a partir das metáforas que foram selecionadas nos pré-testes, p. ex.: “Irene é um furacão”, construída a partir de “Algumas mulheres são furacões” (cf. Tabela 4); no tipo 2: expressões de significado literal do mesmo tipo (X é um Y), p. ex.: “Irene é um furacão”, construída a partir do que se apurou sobre o furacão Irene, que assolou o Caribe e os EUA em 2011; e no tipo 3: declarações de inclusão em classe do mesmo tipo (X é um Y), p. ex.: “Irene é uma jovem”, categorizando o tópico da expressão. Nos contextos do tipo 1, as expressões “X é um Y” demandavam interpretações metafóricas; nos do tipo 2, as mesmas expressões (X é um Y) pediam interpretações literais; e nos do tipo 3, “X é um Y” eram usuais declarações literais de inclusão em classe. Os fragmentos críticos das expressões – “um Y” em todas – metafóricas, literais e declarações literais de inclusão, foram balanceados em número de sílabas, 4; e em tipo e número de número de constituintes, “Det + N” (nos exemplos, respectivamente, “um furacão”, “um furacão”, “uma jovem”). E, ao final da leitura de cada um dos parágrafos, os participantes responderam a uma pergunta acerca do conteúdo lido, pressionando os botões correspondentes a SIM ou Não na caixa de botões acoplada ao dispositivo de exibição dos estímulos. As barras nos exemplos abaixo limitam, no papel, os 11 fragmentos apresentados em cada uma das janelas de leitura, na cadência imposta por cada um dos participantes ao realizar a leitura dos parágrafos.

Contexto do tipo 1: demandando interpretação metafórica

Com muita frequência, / mulheres brasileiras / trabalham fora, / têm filhos / e estudam. / Irene é / um furacão / que empolga / todo mundo / no trabalho, em casa / e na faculdade.

Contexto do tipo 2: demandando interpretação literal

Com muita frequência, / fenômenos naturais / recebem nomes / de gente / dos mais comuns. / Irene é / um furacão / que assolou / o Caribe / e o leste dos EUA / em 2011.

Contexto do tipo 3: demandando interpretação de declaração de inclusão em classe

Com muita frequência, / mulheres brasileiras / trabalham fora, / têm filhos / e estudam. / Irene é / uma jovem / que empolga / todo mundo / no trabalho, em casa / e na faculdade.

Os estímulos foram distribuídos (*within-subjects*, contrabalanceados em *latin square*) de modo que os participantes leram parágrafos em todas as 3 condições (lendo 3 estímulos por condição e jamais lendo o mesmo estímulo em mais de uma condição). Foram intercalados entre os 9 parágrafos-alvo outros 18 parágrafos, distrativos e a ordem de apresentação do total de 27 aleatorizada a cada sessão com cada um dos participantes. As perguntas após cada um dos parágrafos-alvo demandaram dos participantes, aproximadamente, tantas respostas SIM quantas NÃO.

Procedimentos: Em sessões individuais, cada um dos participantes leu os estímulos, fragmento a fragmento, apresentados, sob o seu comando, na tela de um microcomputador Apple, pressionando “o botão de passar a tela” de uma caixa de botões acoplada ao dispositivo. Para isso, a apresentação dos estímulos obedeceu a um protocolo construído em PsyScope, de modo a gravar os tempos de leitura e as respostas dos participantes. Antes de cada sessão, garantiu-se ao participante um breve treinamento, constando da leitura de 4 estímulos, com o fim de familiarizá-lo com o manuseio do equipamento e a rotina do experimento. Aos estudantes dos cursos de graduação foi oferecido um certificado de participação em atividade complementar aos estudos que eles(as) estivessem desenvolvendo. E, ao final de todas as sessões, os participantes puderam conhecer os objetivos do experimento e ter esclarecidas dúvidas que eventualmente tivessem.

Hipótese: De acordo com o modelo teórico de *Class-inclusion* (GLUCKSBERG; KEYSAR, 1990), as metáforas do PB muito familiares altamente convencionalizadas, e *very high-apt* selecionadas nos *norming studies* realizados devem ser processadas diretamente, sem a intermediação da derivação do significado literal, tal qual o são declarações literais de inclusão em classe. Alternativamente, de acordo com o chamado Modelo Pragmático Padrão, as mesmas metáforas devem ser processadas indiretamente, a partir da derivação do significado literal da expressão e sua rejeição no contexto do enunciado.

Predições: Consequentemente à hipótese de pesquisa, de processamento direto de metáforas muito familiares, altamente convencionalizadas e *high-apt*, não se esperam diferenças significativas entre os tempos de leitura dos veículos dessas metáforas e os dos SNs correspondentes, tanto nas expressões equivalentes de significado literal quanto nas declarações literais de inclusão em classe. Prevalecendo o processamento indireto, devem-se observar tempos médios de leitura dos veículos das metáforas significativamente maiores do que os dos mesmos SNs-alvo de interpretação literal e os SNs-alvo nas declarações literais de inclusão em classe.

4.1 Análise dos Resultados do experimento de leitura automonitorada

Adotamos, aqui, o rigor metodológico que, nas críticas de Janus e Bever (1985), falta à literatura de até então, tanto ao sustentar quanto ao refutar o processamento indireto da metáfora que segue o Modelo Pragmático Padrão. Para tanto, dividimos os estímulos em sintagmas e medimos os tempos de leitura (*RTs*) sobre o *locus* do “Problema do Reconhecimento” do significado metafórico, os veículos das metáforas. Mas, diferentemente de Janus e Bever (1985), cujos achados sugerem que metáforas novas são processadas indiretamente, a partir da derivação do significado literal da expressão, obtivemos resultados que oferecem sustentação à hipótese do processamento direto de metáforas convencionalizadas, familiares e *high-apt*. De fato, Janus e Bever (1985, p. 485) admitem a possibilidade de processamento direto, do que eles chamam de *frozen metaphors* e aparentemente não foram outros os nossos achados no PB.

Contrariando as predições associadas à hipótese de processamento indireto nos termos do Modelo Pragmático Padrão, os participantes do experimento de leitura automonitorada que nós conduzimos não despenderam significativamente mais tempo na leitura dos veículos das metáforas, p. ex., “um furacão”, em “Irene é um furacão”, do que na leitura dos mesmos SNs em expressões de interpretação literal, p. ex., “Irene é um furacão”, referindo o furacão que, em 2011, assolou o Caribe. E nem na leitura dos SNs correspondentes nas declarações literais de inclusão em classe, p. ex.: “Irene é uma jovem”, sugerindo, assim, a validade da hipótese de processamento direto via *Class-inclusion / Dual reference* de Glucksberg e Keysar (1990), como demonstram os *RTs* médios reunidos na Tabela 5.

Tabela 5 – Tempos médios de leitura (*RTs* em ms) dos fragmentos críticos (em destaque) nas condições: Literal, Metafórico e Declaração Literal de Inclusão em Classe (DLIC)

Contextos	<i>RTs</i>	Exemplos (extratos dos parágrafos)
Literal	954,25	[...] fenômenos naturais recebem nomes de gente. Irene é um furacão que assolou o Caribe [...].
Metafórico	925,84	[...] mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é um furacão que empolga todo [...].
DLIC	867,89	[...] mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é uma jovem que empolga todo [...].

Em nenhuma das 3 condições, Literal, Metafóra e Declaração Literal de Inclusão em Classe (DLIC), os *RTs* dos fragmentos críticos apresentaram distribuição normal (na condição Literal, Estatística Kolmogorov-Smirnov=0,14, $p < ,05$; na condição Metafóra, Kolmogorov-Smirnov=0,12, $p < ,05$; e na condição DLIC, Kolmogorov-Smirnov=0,12, $p < ,05$). Além disso, não se mostraram diferenças significativas nas Comparações Múltiplas entre as condições, DLIC-Literal, DLIC-metafóra, Literal-Metafóra na análise estatística por sujeito (Kruskal-Wallis $X^2(2)=2,95$, $p > ,05$) e na análise por item (Kruskal-Wallis $X^2(2)=3,55$, $p > ,05$). O Poder do Teste utilizado foi estimado em 70%, capaz de avaliar diferenças de, no mínimo, 200 milissegundos (ms) entre os *RTs* dos fragmentos críticos das condições, diferenças como as que Janus e Bever (1985) relataram e que aqui não apareceram.¹²

¹² A natureza epistemológica de achados como os nossos, que se baseiam na falta de diferenças estatisticamente significativas entre as condições experimentais, tem sido questionada, tradicionalmente, nos termos do chamado *argumentum ad ignorantiam* (em inglês, *the absence of evidence is not evidence of the absence*). A evidência de ausência, na tradição que remonta ao ceticismo de Sextus Empiricus (século II), deve ser considerada falaciosa, como ausência de evidência que apenas pode não ter sido encontrada. Entretanto, cresce em importância a visão que não contempla com a pecha de falácia alguns casos em que uma proposição é presumida verdadeira simplesmente porque não pode ser provada falsa, ou é falsa porque não pode ser provada verdadeira (WALTON, 1992, p. 381-4). Admitem-se entre tais casos circunstâncias em que evidências conclusivas não se mostram apesar de perseguidas; e se presume verdadeira ou falsa a proposição que expressa a falta desses resultados, ainda que provisoriamente, tendo em vista a possibilidade de futuras refutações (p. 383). Sendo assim, alegamos

As respostas dos participantes às perguntas sobre o conteúdo dos parágrafos atestaram, satisfatoriamente, o envolvimento dos indivíduos na realização da tarefa proposta. Considerando a sua totalidade (66), observou-se esmagadora maioria (94%) de respostas corretas. Ou ainda, aproximadamente 99% dos participantes responderam corretamente a 90% das perguntas, permitindo que não se atribuam os resultados à falta de atenção dos leitores ou a dificuldades de compreensão dos estímulos.

A diferença estatisticamente não significativa (apenas 28,4ms) entre o tempo médio de leitura (*RT* médio) dos veículos das metáforas e o dos mesmos SNs interpretados literalmente é sugestiva de que os significados metafóricos familiares, *high-apt* e convencionalizados (assim julgados nas respostas *off-line* dos participantes dos *norming studies*) acham-se, de fato, “disponíveis” ao se dar a estruturação *on-line* das expressões predicativas e sua interpretação em contextos apropriados. Em relação a essa diferença (estatisticamente não significativa, de 28ms, repetimos), pode-se indagar se causaram aos participantes alguma “estranheza” e dificuldade de interpretar as expressões literais que, nos estímulos utilizados, atribuíam “nomes de gente” a entes não humanos, animados ou inanimados, *e.g.*, “Irene é um furacão” no parágrafo “[...] fenômenos naturais recebem nomes de gente. Irene é um furacão que assolou o Caribe [...]” (cf. Apêndice, para os demais). Esse mesmo *RT* médio dos SNs de interpretação literal também se afastou (86,36ms, diferença essa que também não se revelou estatisticamente significativa) do *RT* médio de leitura dos SNs das declarações de inclusão em classe, *p. ex.*, “Irene é uma jovem”, que são do repertório dos falantes do PB enquanto conhecimento semântico consolidado. Se, então, algumas expressões literais causaram “estranhamento” (e expressões literais não

ser esse o caso das diferenças entre os *RTs* dos segmentos críticos nas condições (como Hipótese Alternativa) que não se revelaram no experimento que conduzimos, ensejando a defesa – ao menos presuntiva, nos termos de Walton – da Hipótese de Pesquisa. Em reforço à presunção de que não são falaciosos os achados experimentais que obtivemos, aduzimos que, se, por um lado, a ausência de diferenças significativas entre os *RTs* dos segmentos críticos nas 3 condições pode ser parafraseada em termos da não rejeição da Hipótese Nula, por outro, o Poder do Teste utilizado no tratamento estatístico dos dados, estimado em 70%, seria capaz de surpreender diferenças de, no mínimo, 200ms como aquelas que Janus e Bever (1985) relataram no processamento de metáforas novas, e, portanto, responder ao desafio de se estar cometendo, aqui, o erro do tipo II, de não rejeitar H_0 falsa. (COOLICAN, 2014, p. 427-8).

deveriam causar qualquer um), cabe alegar que o efeito ao confrontar, p. ex., “Fernanda é uma lesma”, no parágrafo “Em certas culturas, animais exóticos são adotados e ganham apelidos. Fernanda é uma lesma que habita um viveiro na casa de criadores nas Filipinas”, deve-se, em tais casos, à “infelicidade” pragmática da declaração literal e não à violação das condições de verdade da sentença (cf. MATTHEWSON, 2004, p. 409).¹³ É verdade que, assim, se estaria estendendo a exemplares dos chamados enunciados “consta(ta)tivos” problemas, em princípio, inerentes aos “performativos”, mas não sem algum aval da literatura. Ao menos para Gerken (1994, p. 78), ao “descrever”, “explicar” e “teorizar” também nos comprometemos em atividades performativas, o que equivale a afirmar que seria mais adequado considerar “constativo” ou “descritivo” como caso especial de “performativo”. Em Petrey (1990), é possível verificar que até Austin (1962) reconheceu que a distinção (originalmente proposta por ele, Austin) era demasiado forte, inclinando-se por sua reavaliação nos termos acima. Em Grandesso (2006), é possível tomar contato com a literatura em língua portuguesa que contempla as visões aqui expostas.

No que diz respeito à hipótese do processamento direto de metáforas, voltamos aqui ao que importa destacar: que o *RT* médio de leitura dos veículos das metáforas não diferiu significativamente do *RT* médio dos mesmos SNs interpretados literalmente. E que, também, não diferiu significativamente do *RT* médio dos SNs correspondentes nas declarações literais de inclusão em classe (57,95ms), o que enseja a defesa da posição Glucksberg e Keysar (1990), de que metáforas são o que aparentam ser: declarações de inclusão em classe, categorias metafóricas que recebem os nomes dos veículos, que delas são membros prototípicos. A não observância nos dados do PB aqui relatados de diferenças significativas entre os *RTs* dos fragmentos críticos das condições sugere que metáforas familiares, *high-apt* e convencionalizadas sejam processadas como declarações de inclusão em classe, conforme o modelo de *Class-inclusion* (GLUCKSBERG; KEYSAR, 1990).

¹³ Agradecemos à Dra. Luciana Sanchez Mendes, nossa colega no GEPEX – UFF, a sugestão de encaminhar na direção seguida acima a avaliação da instigante (embora estatisticamente marginal) diferença entre os tempos de leitura de SNs de interpretação literal e declarações literais de inclusão em classe.

A ausência de diferenças significativas entre os *RTs* que obtivemos poderia ser atribuída, talvez, à constituição da amostra. Entre os 66 participantes do experimento, não havia, de fato, equilíbrio na distribuição entre homens (1/3) e mulheres (2/3); os indivíduos selecionados eram habitantes de diferentes cidades (duas, embora no mesmo estado); e cursavam ou haviam concluído os seus estudos em instituições de Ensino Superior que têm perfis diferentes uma, de ensino público, outra, instituição privada; eram estudantes de Graduação, incompleta (59) e completa (1) e com Especialização (6); e oriundos de cursos diversos.

Certamente, não se pretendeu ignorar o fato de que a compreensão de metáforas por indivíduos inseridos em contextos socioculturais diferentes pode diferir consideravelmente (para uma iniciação aos estudos de aspectos culturais envolvidos na compreensão de metáforas, vejam-se, por exemplo: Gibbs Jr. (2008) e Ortony (1993 [1979]) e referências por eles recomendadas). Tanto é assim que, na primeira fase da pesquisa, foram colhidos julgamentos a respeito de familiaridade, *aptness* e convencionalidade das expressões que deram origem aos estímulos utilizados no experimento de leitura automonitorada.

Ainda assim, passamos, a partir deste ponto, a sistematizar os achados do experimento de leitura automonitorada à luz do que se pode chamar de “uniformização” da amostra global. Ao restringir a amostra aos participantes que deram respostas corretas a todas as perguntas (100% em vez de 94%), e aos que eram estudantes da UFF, cursando o segundo período de Letras – chegou-se a 38 indivíduos cuja média de idade (antes, de 27,3 anos e, agora, 26,6 anos) alterou-se em menos de 1 ano; e a um subconjunto cuja proporção entre homens (25%) e mulheres (75%), reflete o público dos cursos de Letras no país. Considerando, então, a amostra “uniformizada”, não se acham motivos de questionamentos dos resultados anteriormente mostrados (cf. Tabela 6), que não se mostraram divergentes dos anteriormente apontados (na Tabela 5, quando o total de 66 participantes foi considerado) e, portanto, não justificam outras conclusões.

Tabela 6 – Tempos médios de leitura (*RTs*) dos fragmentos críticos nas condições: Literal, Metáfora e Declaração Literal de Inclusão em Classe (em ms), com 66 e 38 participantes

	Literal	Metafórico	Controle
66 participantes	954,25	925,84	867,89
38 participantes	1008,02	968,11	930,11

Consequentemente, de uma maneira ou de outra, não se sustenta, nos resultados do experimento de leitura automonitorada realizado, a hipótese do processamento indireto de metáforas familiares, *high-apt* e convencionalizadas do PB. Os *RTs* médios não diferiram significativamente, sugerindo, assim, que os participantes, tanto os 66 da amostra original quanto os 38 do subconjunto dos que pertenciam a um grupo geográfico e socialmente mais homogêneo, compreenderam, sem dificuldades adicionais, em relação às mesmas expressões mesmas literais e declarações literais de inclusão em classe, as metáforas inseridas em contextos apropriados. Ao interpretar o significado metafórico de, p. ex., “Irene é um furacão” (porque aparentemente Irene agita o ambiente social, por onde quer que ela passe), os falantes nativos do PB da amostra selecionada não encontraram dificuldades, tanto quanto não encontraram em relação a interpretar literalmente “Irene é um furacão” (que assolou o Caribe). Bem como não encontraram dificuldades para interpretar o significado metafórico da mesma expressão maiores do que para interpretar “Irene é uma jovem”, declaração literal de inclusão em classe (incluindo Irene na categoria das jovens), permitindo-nos inferir, como querem Glucksberg e cols. (GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998; GLUCKSBERG, 2003), que compreenderam as metáforas tal qual compreenderam as declarações literais de inclusão em classe.

Figura 3 – Propriedades associadas ao significado metafórico M-furacão de “furacão” no nível supraordenado e ao significado literal L-furacão “furacão” no nível básico da categoria

Nível Supraordenado da Categoria:	
propriedades M- furacão:	Muito impetuoso(a) Avassalador(a) Muito agitado(a) Muito Forte
Nível básico da Categoria:	
propriedades L- furacão:	Muito impetuoso(a) Avassalador(a) Muito agitado(a) Muito Forte Destruidor(a) Muito veloz Devastador(a) Tumultuador(a)

Assim como Glucksberg (1998, p. 41) estrutura, hierarquicamente, em 2 níveis de referência, em 2 níveis de referência, um básico e outro supraordenado ou metafórico, para dar conta de *my lawyer is a shark* (“meu advogado é um tubarão”), a categoria metafórica “furacão” acha-se ilustrada na Figura 3.

De acordo com o modelo *Class-inclusion / Dual reference*, diante da expressão “Irene é um furacão”, o ouvinte / leitor acessa prontamente o significado metafórico do veículo da metáfora, “furacão” – convencionalizado (no PB, até dicionarizado. No modelo de Glucksberg e Keysar (1990) de processamento direto de metáforas por categorização, metáforas como, p. ex., “Irene é um furacão”, seriam compreendidas como são compreendidas as declarações literais de inclusão em classe, às custas da inclusão de tópicos de categorias preexistentes, criadas a partir de veículos convencionalizados.

5 Conclusões

Os resultados dos experimentos realizados sustentam a hipótese defendida por Glucksberg e Keysar (1990), de processamento direto de metáforas familiares, *high-apt* e cujos veículos se acham convencionalizados, conforme o modelo *Class-inclusion*. Foram medidos os RTs sobre os fragmentos críticos das expressões, encaixadas em contextos apropriados, e não se observou diferença significativa entre

os *RTs* dos veículos das metáforas e das expressões literais equivalentes. Tampouco se observou diferença significativa entre os *RTs* médios dos veículos das metáforas e de declarações literais de inclusão em classe, o sugere que o processamento das metáforas em questão deu-se conforme o modelo *Class-inclusion / Dual reference* (GLUCKSBERG; KEYSAR, 1990).

Importa assinalar que aqui não se cogitou estender a ideia de processamento direto via *Class-inclusion* a categorias que não as preexistentes, como ambiciona Glucksberg (2008). Para Glucksberg (2008), a convencionalidade de “veículos” decorre necessariamente do seu uso em expressões bem construídas. E, sendo assim, metáforas *high-apt*, ainda que novas, seriam interpretadas como declarações de inclusão em classe. No futuro, tal hipótese poderá (e deve) ser devidamente testada.

Apesar do ineditismo, o presente estudo não ultrapassa os limites que, desde o início, se impôs: o de avaliar no PB o processamento de metáforas convencionalizadas, familiares e *high-apt*. Fica ainda à espera do interesse de pesquisas que ainda hão de vir, o estudo do processamento de metáforas do PB com o auxílio do rastreamento de movimentos oculares na compreensão de metáforas e, é claro, a extração de potenciais evocados durante a audição / leitura de metáforas do PB e investigações com base em neuroimagem.

Agradecimentos

Agradecemos aos pares anônimos pelas indicações de aspectos do texto que mereceram reformulação. Ao Parecerista 1, no trato da transição para os objetivos, hipótese e predições do experimento de leitura auto-monitorada. Ao Parecerista 2, por pleitear (sem exigir) o uso de equivalentes em português de termos em inglês que se acham consagrados na literatura internacional, o que virá, cada vez mais, à medida que o campo de processamento de metáfora se consolidar entre nós. E, finalmente, ao organizador do volume por nos impelir à necessária discussão (veja-se a nota 12) dos achados que relatamos, tendo em vista as ideias em torno do chamado *argumentum ad ignorantiam*.

Referências

AULETE, Caldas. *Aulete Digital*: o dicionário da língua portuguesa na internet. Versão online. Disponível em: <<http://www.aulete.com.br/>>. Acesso em: 16 set. 2015.

AUSTIN, J.L. *How to do things with words*. Cambridge, Mass.: Harvard University Press, 1962.

BLASKO, D.; CONNINE, C.M. Effects of familiarity and aptness on metaphor processing. *The Journal of Experimental Psychology: Learning, Memory, and Cognition*, Washington, v. 19, p. 295-308, 1993. <https://doi.org/10.1037/0278-7393.19.2.295>.

BOWDLE, B. F.; GENTNER, D. The career of metaphor. *Psychological Review*, American Psychological Association, v. 112, n. 1, p. 193-216, 2005. <https://doi.org/10.1037/0033-295x.112.1.193>.

CHIAPPE, D.; KENNEDY, J. M.; SMYKOWSKI, T. Reversibility, aptness, and the conventionality of metaphors and similes. *Metaphor and Symbol*, v. 18, n. 2, p. 85-105, 2003. https://doi.org/10.1207/S15327868MS1802_2.

CLARK, H. H.; LUCY, P. Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, Elsevier B. V., n. 14, p. 56-72, 1975.

COOLICAN, H. *Research methods and statistics in psychology*. 6th ed. London; New York: Psychology Press; Taylor & Francis Group, 2014.

CRAIG, Colette G. (Ed.). *Noun classes and categorization*: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983. John Benjamins Publishing Company, 1986.

DENNY, J. Peter. The semantic role of noun classifiers. *Noun classes and categorization*, p. 297-308, 1986.

DULCINATI, G.; MAZZARELLA, D.; POUSSCOULOUS, N.; RODD, J. Processing metaphor: The role of conventionality, familiarity and dominance. *UCL Working Papers in Linguistics*, London, University College London, v. 26, p. 72-88. 2014.

FERREIRA, Aurélio Buarque de Hollanda. *Miniaurélio*: o Minidicionário da Língua Portuguesa. 6. ed. ver. ampl. Curitiba: Positivo, 2004.

FERREIRA, Aurélio Buarque de Hollanda. *Novo Aurélio século XXI: o dicionário da língua portuguesa*. 3. ed. ver. ampl. Rio de Janeiro: Nova Fronteira, 1999.

FRASER, B. The interpretation of novel metaphors. In: ORTONY, A. (Ed.). *Metaphor and thought*. Cambridge: Cambridge University Press, p. 329-341, 1993. <https://doi.org/10.1017/CBO9781139173865.017>.

GARRET, J. *Aristotle on metaphor*. 2007. Disponível em: <<http://people.wku.edu/jan.garrett/401s07/arismet.htm>>. Acesso em: 6 dez. 2016.

GERKEN, L. A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, Cambridge, Cambridge University Press, v. 21, p. 565-584, 1994. <https://doi.org/10.1017/s0305000900009466>.

GIBBS JR, R. W. Contextual effects in understanding indirect requests. *Discourse Processes*, v. 2, p. 1-10, 1979. <https://doi.org/10.1080/01638537909544450>.

GIBBS JR, R. W. *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, 2008. <https://doi.org/10.1017/CBO9780511816802>.

GIBBS JR, R. W. Your wish is my command: Convention and context in interpreting indirect requests. *Journal of Verbal Learning and Verbal Behavior*, Elsevier B. V., v. 20, p. 431-444, 1981.

GLUCKSBERG, S. How metaphors create categories – quickly. In: GIBBS JR, R. W. *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, p. 67-83, 2008. <https://doi.org/10.1017/CBO9780511816802>.

GLUCKSBERG, S. The Psycholinguistics of metaphor. *Trends in Cognitive Sciences*, Elsevier, v. 7, n. 2, p. 92-96, 2003. [https://doi.org/10.1016/S1364-6613\(02\)00040-2](https://doi.org/10.1016/S1364-6613(02)00040-2).

GLUCKSBERG, S. Understanding metaphors. *Current Directions in Psychological Science*, Cambridge, Cambridge University Press, v. 7, n. 2, p. 39-43, 1998.

GLUCKSBERG, S.; GILDEA, P.; BOOKIN, H. B. On understanding nonliteral speech: can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, Elsevier B. V., v. 21, p. 85-98, 1982.

GLUCKSBERG, S.; KEYSAR, B. Understanding Metaphorical Comparisons: Beyond Similarity. *Psychological Review*, American Psychological Association, v. 97, n. 1, p. 3-18, 1990. [https://doi.org/10.1016/S1364-6613\(02\)00040-2](https://doi.org/10.1016/S1364-6613(02)00040-2).

GRANDESSO, M. *Sobre a reconstrução do significado: uma análise epistemológica e hermenêutica da prática clínica*. 2. ed. São Paulo: Casa do Psicólogo, 2006.

GRICE, H. P. Logic and Conversation. In: COLE, P.; MORGAN, J. L. (Ed.). *Syntax and Semantics*. New York: Academic Press, 1975. [v. 3: Speech Acts, p. 41-58].

HAGE, P.; MILLER, W. R. 'Eagle' = 'bird': A note on the structure and evolution of Shoshoni ethnoornithological nomenclature. *American Ethnologist*, 3, p. 481-488, 1976. <https://doi.org/10.1525/ae.1976.3.3.02a00060>.

HARRIS, R. J. Comprehension of metaphors: A test of the two-stage processing model. *Bulletin of the Psychonomic Society*, v. 8, n. 4, p. 312-314, 1976. <https://doi.org/10.3758/BF03335150>.

HOUAISS, A.; VILLAR, M. S. *Dicionário Houaiss da Língua Portuguesa*. 3. ed. Rio de Janeiro: Objetiva, 2009.

JANUS, R. A.; BEVER, T. G. Processing of metaphoric language: an investigation of the three-stage model of metaphor comprehension. *Journal of Psycholinguistic Research*, Springer International Publishing A. G., v. 14, n. 5, p. 473-487, 1985.

JOHNSON, A.T. Comprehension of metaphors and similes: a reaction time study. *Metaphor and Symbolic Activity*, Lawrence Erlbaum Associates, Inc., v. 11, n. 2, p. 145-159, 1996.

JONES, L. L.; ESTES, Z. Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language*, Elsevier, v. 55, n. 1, p. 18-32, 2006. <https://doi.org/10.1016/j.jml.2006.02.004>.

MATTHEWSON, L. On the methodology of semantic fieldwork. *International Journal of American Linguistics*, Chicago, University of Chicago, v. 70, n. 4, p. 369-415, Oct. 2004.

MILLER, G. A. Images and models: similes and metaphors. In: ORTONY, A. (Ed.). *Metaphor and thought*. Cambridge: Cambridge University Press, p. 202-250, 1979.

MORISON, Benjamin. Sextus Empiricus. In: ZALTA, Edward N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Stanford: Spring Edition, 2014. Disponível em: <<https://plato.stanford.edu/archives/spr2014/entries/sextus-empiricus/>>. Acesso em: 1 mar. 2017.

ORTONY, A. *Metaphor and thought*. 2nd edition. New York: Cambridge University Press, 1993 [1979]. <https://doi.org/10.1017/CBO9781139173865>.

ORTONY, Andrew et al. Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of verbal learning and verbal behavior*, v. 17, n. 4, p. 465-477, 1978.

PETREY, S. *Speech Acts and Literary Theory*. New York; London: Routledge, 1990.

RIBEIRO, A. J. C.; RICCI, A.Q.; KENEDY, E. O processamento psicolinguístico da metáfora: um estudo inédito no português brasileiro. In: ENCONTRO NACIONAL DA ANPOLL (ENANPOLL), XXXI, 2016, Campinas. *Poster*. Campinas: UNICAMP, 2016.

RICCI, A. Q. *O processamento psicolinguístico da metáfora: um estudo experimental no PB*. 2016. 75f. Dissertação (Mestrado em Estudos de Linguagem) – Instituto de Letras, Universidade Federal Fluminense, 2016.

SEARLE, J. Metaphor. In: ORTONY, A. *Metaphor and thought*. 2nd edition. New York: Cambridge University Press, 1993. p. 83-111. [1979]

SHINOFF, P. Demjanjuk war-crimes tribunal strikes deep fear among Jews. *San Francisco Examiner*, June 14, p. A8, 1987.

STROOP, J.R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, American Philosophical Association, v. 18, p. 643-662, 1935.

THIBODEAU, P. H.; DURGIN, F. H. Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol*, Francis & Taylor Online, v. 26, p. 206-226, 2011.

TRAGER, G. L. "Cottonwood-Tree", a south-western linguistic trait. *International Journal of American Linguistics*, 9, p. 117-118, 1938. <https://doi.org/10.1017/CBO9781139173865>.

TVERSKY, A. Features of similarity. *Psychological Review*, American Psychological Association, v. 84, p. 327-352, 1977. <https://doi.org/10.1037/0033-295x.84.4.327>.

WALTON, Douglas. Nonfallacious arguments from ignorance. *American Philosophical Quarterly*, v. 29, n. 4, Oct. 1992. Disponível em: <<http://dougwalton.ca/papers%20in%20pdf/92nonfallacious.pdf>>. Acesso em: 1 mar. 2017.

Apêndice: Parágrafos experimentais do experimento cronométrico

Os parágrafos experimentais e respectivas perguntas que se seguem estão organizados na sequência Literal / Metáfora / Declaração Literal de Inclusão em Classe, em cada trinca.

1. Com muita frequência, fenômenos naturais recebem nomes de gente dos mais comuns. Irene é um furacão que assolou o Caribe e o leste dos EUA em 2011.

Com muita frequência, mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é um furacão que empolga todo mundo no trabalho, em casa e na faculdade.

Com muita frequência, mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é uma jovem que empolga todo mundo no trabalho, em casa e na faculdade.

2. Excepcionalmente, móveis italianos têm qualidade e nome no mercado. Bérghamo é um armário que recebe dos clientes ótimas avaliações nos comentários.

Excepcionalmente, segurança de rua porta armas para ser respeitado. Bérghamo é um armário que trabalha com ajuda do porte avantajado e da força bruta.

Excepcionalmente, segurança de rua porta armas para ser respeitado. Bérghamo é um vigia que trabalha com ajuda do porte avantajado e da força bruta.

3. Em certas culturas, animais exóticos são adotados e ganham apelidos. Fernanda é uma lesma que habita um viveiro na casa de criadores nas Filipinas.

Costumeiramente, motoristas idosos guiam devagar sem razão no trânsito. Fernanda é uma lesma que atrasa a chegada no trabalho ou lazer quando usa carro.

Costumeiramente, motoristas idosos guiam devagar sem razão no trânsito. Fernanda é uma chofer que atrasa a chegada no trabalho ou lazer quando usa carro.

4. Em certas famílias, bichos de estimação distanciam-se do padrão esperado. Tereza é uma cobra que só pica raramente e não tem qualquer veneno, segundo o dono.

Em certas famílias sogras problemáticas implicam muito com genros, sem motivos. Tereza é uma cobra que ataca sem aviso o marido da filha em qualquer lugar.

Em certas famílias sogras problemáticas implicam muito com genros, sem motivos. Tereza é uma sogra que ataca sem aviso o marido da filha em qualquer lugar.

5. Em alguns lugares, primatas em extinção ganham proteção de ONGs e pessoas. Samuel é uma fera que resiste na África à extinção dos gorilas até o momento.

Por todo o mundo, guitarristas famosos sobem em palcos ou tocam ao ar livre. Samuel é uma fera que fascina o público da banda de rock Skank há muitos anos.

Por todo o mundo, guitarristas famosos sobem em palcos ou tocam ao ar livre. Samuel é um músico que fascina o público da banda de rock Skank há muitos anos.

6. Em filmes infantis, grupos de felinos formam famílias felizes e saudáveis. Maria é uma gata que alegra as crianças na famosa animação *Os Aristogatas*.

Em qualquer época, meninas sedutoras estão no centro de festas e baladas. Maria é uma gata que encanta os garotos em qualquer reunião de adolescentes.

Em qualquer época, meninas sedutoras estão no centro de festas e baladas. Maria é uma moça que encanta os garotos em qualquer reunião de adolescentes.

7. De tempos em tempos, animais adestrados ficam famosos em filmes de sucesso. Argento é um cavalo que estrela a história *Cavaleiro Solitário* na pele de Silver.

Em alguns lugares, maridos ciumentos podem agredir esposas sem punição. Anderson é um cavalo que recebeu até hoje apenas advertências por bater na mulher.

Em alguns lugares, maridos ciumentos podem agredir esposas sem punição. Anderson é um marido que recebeu até hoje apenas advertências por bater na mulher.

8. Muito raramente, heroínas de filmes são veículos a motor femininos. Rochelle é um avião que atua no sucesso de animação da Disney chamado *Aviões*.

De modo esperado, modelos de beleza são convidadas a posar para fotos. Rochelle é um avião que desfila com sucesso nas capas de revistas e nas passarelas.

De modo esperado, modelos de beleza são convidadas a posar para fotos. Rochelle é uma jovem que desfila com sucesso nas capas de revistas e nas passarelas.

9. Para quase todos, produtos japoneses são tidos como muito bons e duráveis. Tagima é um violão que músicos destacados têm como referência de qualidade.

De modo habitual, garotas de biquíni são criticadas com rigor por mulheres. Tagima é um violão que revela destemor diante dos olhares de todas na praia.

De modo habitual, garotas de biquíni são criticadas com rigor por mulheres. Tagima é uma moça que revela destemor diante dos olhares de todas na praia.

An experimental study on Brazilian Portuguese metaphor processing

Um estudo experimental do processamento de metáforas do português brasileiro

Antonio João Carvalho Ribeiro

Grupo de Estudos e Laboratório em Psicolinguística Experimental – GEPEX
Universidade Federal Fluminense, Niterói, Rio de Janeiro / Brasil
ajcarvalhoribeiro@gmail.com

Adiel Queiroz Ricci

Grupo de Estudos e Laboratório em Psicolinguística Experimental – GEPEX
Universidade Severino Sombra, Vassouras, Rio de Janeiro / Brasil
adielricci@gmail.com

Abstract: An experimental study on the psycholinguistic processing of Brazilian Portuguese attributive metaphors (X is a Y), e.g., “*Irene é um furacão*” (“Irene is a hurricane”), was carried out with the aim of highlighting, from reading times (RTs), the understanding of familiar, high-apt (“well-built”) expressions, the vehicle of which is conventionalized. In the first phase of the research, two norming studies were carried out, aimed at the ranking of attributive metaphors, e.g. “*Algumas mulheres são furacões*” (“Some women are hurricanes”), regarding familiarity, aptness (adaptation), and conventionality. In the second phase of the research, a self-paced, non-cumulative, moving-window reading experiment was conducted, using, for the composition of the stimuli, the metaphors, e.g., “*Irene é um furacão*”, which have reached, in the normative studies of the first phase, ratings of “very familiar”, “very high-apt”, and “highly conventionalized”. Brazilian Portuguese evidence in favor of the direct processing of metaphors was

obtained, as recommended by the Class-inclusion model of Glucksberg and Keysar (1990), since there were no significant differences between the average RTs in the three conditions: “Metaphor”, “Literal” and “Literal Declaration of Class Inclusion.” In contrast to the findings of Janus and Bever (1985), who observed reading times of new metaphors that were significantly longer than those of literal expressions, according to the predictions of the Standard Pragmatic Model of indirect processing.

Keywords: psycholinguistic processing of metaphor; class-inclusion; dual reference; norming studies; figurative language comprehension; Brazilian Portuguese.

Resumo: Um estudo experimental do processamento psicolinguístico de metáforas nominais do português brasileiro (X é um Y), p. ex., “Irene é um furacão”, foi realizado com o objetivo de evidenciar, a partir de tempos de leitura (RTs), a compreensão de expressões familiares, *high-apt* (“bem construídas”) e cujo veículo se acha convencionalizado. Na primeira fase da pesquisa, realizaram-se dois *norming studies* (“estudos normativos”) com vistas ao ranqueamento de metáforas nominais (p. ex. “Algumas mulheres são furacões”) em relação a familiaridade, *aptness* (“adequação”) e convencionalidade. Na segunda fase da pesquisa, um experimento de leitura automonitorada (*self-paced, non-cumulative, moving-window reading*) foi conduzido, recorrendo, para a composição dos estímulos, às metáforas, p. ex., “Irene é um furacão”, que alcançaram, nos estudos normativos da primeira fase, *ratings* (ou “classificações”) de “muito familiares”, “*very high-apt*” e “altamente convencionalizadas”. Evidências do português brasileiro em favor do processamento direto de metáforas foram obtidas, conforme preconiza o modelo de *Class-inclusion*, de Glucksberg e Keysar (1990), pois não se revelaram diferenças significativas entre os RTs médios nas três condições: “metáfora”, “literal” e “declaração literal de inclusão em Classe”, em contraposição aos achados de Janus e Bever (1985), que observaram tempos de leitura de metáforas novas significativamente maiores do que os de expressões literais, conforme as predições do Modelo Pragmático Padrão de processamento indireto.

Palavras-chave: processamento psicolinguístico da metáfora; *class-inclusion*; *dual reference*; *norming studies*; compreensão de linguagem figurada; português brasileiro.

Received on December 8, 2017.

Approved on March 16, 2017.

1 Introduction

This article intends to include Brazilian Portuguese (BP) in the literature about understanding metaphors, polarized, according to Glucksberg (2003), between indirect processing, whose assumptions have proven testable hypothesis, and direct processing models, based on empirical evidence of interference caused by “Familiarity”, “Aptness”, and “Conventionality” variables. To achieve this, we report on the results for the research in which we have been involved for the past two years, investigating the understanding of metaphors that have been conventionalized in BP, which are suggestive of direct processing, according to Glucksberg and Keysar’s Class-inclusion model (1990).¹

During the initial research phase, two pre-tests were performed – norming studies, in which participants ranked metaphorical expression, such as, “*Algumas mulheres são furacões*” (“Some women are hurricanes”), according to Familiarity, Aptness, and Conventionality.² In linguistics literature, the instruments conceived to evaluate metaphorical expressions are called norming studies (BLASKO; CONNINE, 1993; BOWDLE; GENTNER, 2005; DULCINATI et al., 2014), or norming pre-tests (JONES; ESTES, 2006), or even rating experiments (THIBODEAU; DURGIN, 2011), or, simply, rating tasks (HARRIS, 1976). Rankings were designed to select expressions, among those that received the highest ratings in both, as guiding the construction of stimuli for the self-monitored reading experiment conducted in the subsequent phase. For instance, high ratings assigned to “*Algumas mulheres são furacões*” (“Some women are hurricanes”) were the basis for building the stimulus “*Irene é um furacão*” (“Irene is a hurricane”), which, included in appropriate contexts (paragraphs), would enable the metaphorical interpretation (Irene being a hectic woman) or a literal one (Irene being the hurricane that hit the Caribbean and the USA in 2011).

As the majority of experiments described in the literature focus on this form (DULCINATI et al., 2014, p. 72), we opted to use nominal

¹ The research in question, advised by the first author, was conducted by the latter author, as part of the requirements for obtaining the title of Master in Linguistics (RICCI, 2016). Approval by USS – RJ CEP, CAAE 50572215.80000.5290, report 1 412 627, on Feb. 22, 2016.

² Results of the aforementioned norming studies were presented during the XXXI ENANPOLL (RIBEIRO et al., 2016)

metaphors – such as X is a Y – in which X is the topic and Y is the vehicle; in the examples above, “*Algumas mulheres*” (“Some women”) and “Irene” / “*furacões*” (“hurricanes”) and “*furacão*” (“hurricane”), respectively.

Finally, the decision to propose a self-monitored reading task involving BP expressions in several contexts was designed to overcome the methodological frailties identified by Janus and Bever (1985) in the experiments conducted up to then - supporting (cf. CLARK; LUCY, 1975; GIBBS JR., 1979, 1981) or refuting (cf. ORTONY et al., 1978; GLUCKSBERG; GILDEA; BOOKIN, 1982) what is conventionally called the Standard Pragmatic Model of indirect metaphor processing (cf. SEARLE, 1993 [1979]; GRICE, 1975) – in which context is not provided, nor is reading time measured after the stimuli have been presented. Janus and Bever (1985) measured reading time spent on critical fragments of contextualized expressions, that is, of nominal syntagmas that contain the vehicle – locus of “the Problem to Recognize” metaphorical meaning – and noticed that “new” metaphors took, on average, significantly longer reading times than did equivalent expressions with literal meanings. Therefore, the literature on the topic has produced a gap that this BP study intends to fill: by adopting the key approach used in the experiment conducted by Janus and Bever (1985), it offers chronometric findings in favor of the direct processing of familiar, high-apt metaphors, whose vehicle is conventionalized.

2 Literature review

Traditionally, echoing the Aristotelian view of metaphor as a deviation of ordinary language (GARRET, 2007), metaphor interpretation is considered optional (GLUCKSBERG, 2003; BOWDLE; GENTNER, 2005). According to what is conventionally called the Standard Pragmatic Model (SPM), metaphor processing (as well as that of figurative language in general) takes place in three stages. Initially, the listener/reader derives the literal meaning of the expression. In the second stage, the listener/reader decides if the interpretation is appropriate in the context of the utterance and should be accepted, or not, with the meaning intended by the speaker. If it is implausible, the initial representation is rejected, and the third stage is reached, that of seeking metaphorical interpretation. Thus, the failed initial interpretation of literal meaning works as a “trigger” that sets off the search for a metaphorical meaning that accounts for

the utterance (SEARLE, 1993 [1979], p. 89-90). Therefore, nominal metaphors, such as, for instance, “Some jobs are jails” (example obtained from GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998, 2003) would be recognized as assertions that do not meet the maxim of quality (GRICE, 1975), and one way of understanding them would be to convert them into similes, “Some jobs are like jails”, true literal statements, since any two given things may be similar in numerous ways.

However, Glucksberg, Gildea, and Bookin (1982) question the priority of literal meaning/optionality of metaphorical interpretation, an assumption in SPM, based on evidence obtained from a series of sentence check experiments inspired by Stroop’s test (1935). Stroop (1935) illustrated the competition among attentional processes that enables one to capture the automatism of the initial reading processing phases, translated into errors or longer reaction times in answering questions about the color in which a word (e.g., red) is written, when it is printed in a different color (e.g., green). Glucksberg, Gildea, and Bookin (1982.) argue that participants took longer to consider the metaphors they were presented, e.g., “some surgeons are butchers”, as false, than it took them to consider “jumbled” metaphors, e.g. “some jobs are butchers”, as false, since the metaphorical meanings available in the metaphors, rather than the jumbled metaphors, interfered in the evaluation of these expressions’ literal meanings. By resorting to the Stroop effect logic, or to the “compulsion” to read the word when prompted to use the name of the color in which it is printed, Glucksberg, Gildea, and Bookin (1982) refute the SPM model, according to which metaphorical interpretations are only considered when literal meanings prove themselves unmanageable, supporting the theory that metaphorical interpretations may not be inhibited, or, in their own words, that the metaphorical meaning may not be ignored.

According to Glucksberg and his collaborators (GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998; GLUCKSBERG, 2003), the different theories (cf. TVERSKY, 1977; ORTONY, 1993 [1979]) that treat metaphors as implied comparisons fail because, in literal comparisons, the equation “X is like Y” is bidirectional, e.g.: “copper is like tin” or “tin is like copper”, “coffee is like tea” or “tea is like coffee”, while, in metaphors, the elements may not be switched, e.g.: “some jobs are jails”, but not “*some jails are jobs”, “sermons are sleeping pills”, but not “*sleeping pills are sermons”. The apparent reversibility of metaphor elements produces a new expression, e.g.: “some surgeons are butchers”

and “some butchers are surgeons” are different in the foundations that sustain the construction of each one of them.

According to Glucksberg and his collaborators (GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998), metaphors are what they seem to be: class inclusion statements, as well as literal assertions of category inclusion. The metaphor “sermons are sleeping pills”, for instance, behaves as a statement of inclusion of “sermons” in the metaphorical category of “sleeping pills”, as well as the assertion “a tree is a plant” includes “tree” in the “plant” class (GLUCKSBERG; KEYSAR, 1990, p. 12). According to them, evidence that, in nominal metaphors, “X is Y” expresses the class inclusion relation, rather than an equality relation, is the possibility of easily paraphrasing metaphorical comparisons in metaphors (in the paradigmatic form), e.g.: “my job is like a jail” → “my job is a jail”; whereas it is impossible to paraphrase literal comparisons in true statements, e.g.: “bees are like hornets” → “*bees are hornets” (GLUCKSBERG; KEYSAR, 1990, p. 7-11) or “my lawyer was like a shark” → “my lawyer was a shark”, but not “barracudas are like sharks” → “*barracudas are sharks” (GLUCKSBERG, 1998, p. 41).

According to Glucksberg and Keysar’s theoretical model (1990) – Class-inclusion – metaphors are understood as class, or category, inclusion statements, in which the metaphorical vehicles have what they call Dual Reference. For instance, “shark” refers to a marine predator (or a category of those) as well as an abstract category of predatory beings in general, to which it lends its name. In “my lawyer is a shark”, the expression vehicle (“shark”) defines a taxonomic relation between marine predator and lawyer, both housed in the abstract category of predatory beings in general, which receives the name of the prototypical member: “shark.” In short, metaphors whose vehicles refer to conventionalized metaphorical categories are processed directly, as are literal category inclusion assertions, in the terms of the class-inclusion/dual reference model, in which the latter provides explanation both for the “metaphor’s non-reversibility” and for the possibility of paraphrasing metaphorical comparisons in statements, which enables one to distinguish metaphors from comparisons (GLUCKSBERG, 2003, p. 95). Understanding these facts, Glucksberg (1998, 2003) rejects the hypothesis that metaphors would be processed based on their conversion into similes, supported by Johnson’s empirical findings (1996), that metaphors in the form of a statement, e.g.: “my lawyer is a shark” were understood more quickly

than similes, “my lawyer is like a shark”, despite the latter having an additional word (“like”) when compared to the first.

By legitimating the possibility that the prototypical member designates the metaphorical category, Glucksberg and Keysar (1990, p. 8) argue that several languages use the strategy to label supraordinate categories routinely. American Sign Language uses signs for basic furniture items, “chair”, “table”, “bed”, to refer to the “furniture” category. Burmese, spoken in Southeast Asia (Burma, Thailand, Malaysia, and Singapore), resorts to the double function of names that refer to highly prototypical objects and classes (DENNY, 1986; CRAIG, 1986 *apud* GLUCKSBERG; KEYSAR, 1990). Even non-classifying languages, such as Hebrew and English, provide examples of using prototypical names to designate categories, such as the last name of someone who was accused of war crimes, Demjanjuk, who, in Israel and in North-American articles and newspapers, became the word to identify a “common individual capable of unspeakable acts” (SHINOFF, 1987 *apud* GLUCKSBERG; KEYSAR, 1990) – although John Demjanjuk, was not found guilty, ultimately, of the crimes attributed to the sadistic guard at the Treblinka concentration camp, in Poland, nicknamed “Ivan, the Terrible”. In a large number of indigenous languages from the southwestern regions of the United States, there are plenty of examples of the occasional names of prototypical members to label the categories in which they are included. In the Hopi language, the name of the most abundant species, cottonwood, may designate all trees called deciduous trees, which lose their leaves completely during the fall and winter (TRAGER, 1938 *apud* GLUCKSBERG; KEYSAR, 1990). In the Shoshoni language, the word for “eagle” is also used for large birds in general (HAGE; MILLER, 1976 *apud* GLUCKSBERG; KEYSAR, 1990). According to Glucksberg and Keysar (1990), the underlying principle for the dual reference examples above – using the prototypical member to refer to the category – explains what happens in metaphors: the vehicle is used to refer to the metaphorical, abstract category, in addition to referring to the actual being. There are cases in which the metaphor vehicle’s dual reference is explicit, such as in the speech of an spectator present during John Demjanjuk’s trial, transcribed by Glucksberg and Keysar (1990, p. 8): I know his name is Demjanjuk, but I don’t know if he is a Demjanjuk. It is also explicit in the example provided by Glucksberg (2003, p. 94): “Cambodia was Vietnam’s Vietnam”, in which the first mention of ‘Vietnam’ refers to the country, whereas the second reference points to the supraordinate abstract category of disastrous military

interventions resulting from the armed conflict in southeast Asia (and vice versa in Portuguese).

Finally, in mapping the metaphor psycholinguistics research field explored thus far, Glucksberg (2003) highlights the evidence in Blasko and Connine's cross-modal priming experiments (1993) (in this case, semantic priming associated with a lexical choice task), in which high familiarity and aptness ratings accelerate the activation of the metaphorical meaning of the metaphor vehicle. Thus, Glucksberg (2003) lists, in addition to the vehicle conventionality, the role played by the Familiarity and Aptness variables in the fast access to the metaphorical meaning, and, therefore, tributary of the direct processing of metaphorical expressions.

3 Experiment 1: Familiarity and Conventionality norming study

Experiment 1 was designed to assess participants' familiarity in relation to metaphors that have been constructed for such a purpose by the researcher. Familiarity is understood here as subjective familiarity, according to Blasko and Connine's definition (1993, p. 305): "the perceived experience with the metaphor."³ Beyond obtaining familiarity ratings for metaphors – classified as familiar and unfamiliar – the conventionality of expressions was estimated, whose figurative meaning of vehicles was collected from prestigious BP dictionaries.⁴ To achieve this, each of the participants chose, among three words, the one that in his/her opinion captured the meaning the expression intended to communicate. Based on the answers, it was determined whether or not the vehicle (Y) had evoked the figurative meaning, or extended meaning recorded in the reference dictionaries, allowing or not allowing one to infer whether this meaning was or was not conventionalized in the sample's individual's repertoire.

³ As opposed to objective familiarity, estimated based on metaphorical frequency, by using the Google search engine with corpus (THIBODEAU; DURGIN, 2011, p. 209).

⁴ AULETE. *Aulete Digital*: o dicionário da língua portuguesa na internet; HOUAISS; VILLAR. *Dicionário Houaiss da Língua Portuguesa*, 3. ed. (2009); FERREIRA. *Novo Aurélio século XXI*: o Dicionário da Língua Portuguesa (1999); and the 6th edition of FERREIRA. *Miniaurélio: o Minidicionário da Língua Portuguesa* (2004).

Participants: 81 students from the Severino Sombra University, in Vassouras, RJ, Brazil, between 18 and 56 years of age, 16 males and 65 females, all native speakers of BP, who had never participated in experiments about metaphor understanding and who had volunteered to participate.

Materials: 84 nominal metaphors (X is a Y) were constructed, by the researcher, with no repetitions in the vehicles (Y) or topics (X). Topics in the expressions were modified by “some”, and both topics (X) and vehicles (Y) are concrete nouns, with the same gender, for instance: “*Alguns carros são abacaxis*” (“Some cars are lemons”). Some topics (X) are exceptions, as they are common gender nouns, and some vehicles are overly common nouns, e.g.: “*Alguns motoristas são lesmas*” (“Some drivers are snails”). Figurative meanings of all vehicles (Y) are recorded in at least two of the reference dictionaries. Understood literally, all constructed expressions are false.

Procedures: On pre-defined days, groups of participants (maximum of 19) were placed in one of the rooms at USS – RJ⁵ Distance Learning Support Center, and, after having signed the Free and Informed Consent Form, they performed the proposed task in, on average, 15 minutes. Each of the participants sat before a 17” LCD screen exhibiting the Google form (Google Chrome on Windows 7) that contained the questions to be answered (each of the screens was connected to a PC with ABNT2 wired keyboard and a wired optical USB mouse with three buttons and scroll). In the heading on each of the form blocks, immediately above the expression they were expected to evaluate (as familiar or unfamiliar), the participants could read the instructions in Portuguese, regarding the procedure to be adopted: “Do the expressions below sound familiar to you?” “Have you ever read or heard them?” And it continued: “Check YES (if it is familiar) or NO (if it is not familiar)”. “After that, if you have heard or read the expression before, choose a word that, in your opinion, captures its meaning.”

Hence, it was possible to estimate the conventionality of the Vehicle (Y) of each of the metaphors, whose figurative meanings were listed among the choices, as well as to assess the quality of the participation of sample components, since the other choices would

⁵ Distance Learning Support Center at the Severino Sombra University in Vassouras, RJ, Brazil.

refer to the expression's literal meaning or to the metaphor's opposite. In distributing multiple choices, we were careful to alternate, within the form, the positions where the options appeared (see excerpt of the form in Figure 1). Answering all questions was mandatory, thus making it impossible for participants to leave any questions blank when they submitted the form on the internet. At the end of each session, participants were given access to the purposes of the experiment and were able to clarify any potential doubts.

Figure 1 – Excerpt from the Familiarity form

FAMILIARIDADE
 As expressões abaixo lhe são familiares? Você já as leu ou ouviu?
 Marque SIM (se for familiar) ou NÃO (se não for familiar).
 Depois, tendo, ou não, ouvido ou lido a expressão, escolha uma palavra que você acredita que capture o seu significado.

1. Alguns carros são abacaxis. *
 (Esta expressão lhe é familiar?)

SIM
 NÃO

Agora escolha uma palavra que capture o significado da expressão acima: *

COMPLICAÇÃO
 FACILIDADE
 FRUTO

3.1 Experiment 2: Aptness and Conventionality norming study

Experiment 2 was designed to evaluate the ‘aptness’ of the expressions in the set of 84 metaphors, whose familiarity was assessed in Experiment 1. In accordance with the most widely accepted definition of aptness: “the extent to which the statement captures important features of the topic” (CHIAPPE et al., 2003, p. 97),⁶ participants were asked to rate each of the expressions from 1 to 5, evaluating their effectiveness

⁶ Dulcinati et al. (2014, p. 74), warn that this metaphor property has been assigned several different definitions.

in conveying the meaning they intend to convey. In addition, similar to Experiment 1, the expressions' conventionality was also estimated.

Participants: 79 students from Severino Sombra University in Vassouras, RJ, Brazil, 19 to 44 years of age, 28 males and 51 females, all native speakers of BP, who had never participated in experiments about metaphor understanding and who volunteered to participate.

Materials: The same 84 nominal metaphors (X is a Y) provided to participants in Experiment 1.

Procedures: On pre-defined days, groups of participants (maximum of 19) were placed in one of the rooms at the USS – RJ⁷ Distance Learning Support Center, and after having signed the Free and Informed Consent Form, they performed the proposed task in, on average, 15 minutes. Each of the participants sat before a 17" LCD screen exhibiting the Google form (Google Chrome on Windows 7) that contained the questions to be answered (each of the screens connected to a PC with an ABNT2 wired keyboard and a wired optical USB mouse with three buttons and scroll). In the heading on each of the form blocks, immediately above the expression they were expected to evaluate, the participants could read the instructions in Portuguese, regarding the procedure to be adopted: "Rate each of the expressions below from 1 to 5, evaluating their effectiveness in conveying the meaning they intend to convey" and "After that, choose a word that, in your opinion, captures its meaning. Similar to Experiment 1, we were able to estimate the conventionality for each metaphor vehicle (Y), whose figurative meanings were listed among the choices, as well as to assess the quality of the participation of the sample components, as the other choices would refer to the expression's literal meaning or to the metaphor's opposite. In distributing multiple choices, we were careful, within the form, to alternate the positions where the options appeared (see excerpt of the form in Figure 2). Answering all questions was mandatory, thus making it impossible for participants to leave any questions blank when they submit the form on the internet. Similar to Experiment 1, at the end of each session, participants were given access to the purposes of the experiment and were able to clarify any potential doubts.

⁷ At the same facilities where Experiment 1 was held.

Figure 2 – Excerpt from the Aptness form

APTNESS
 Dê uma nota, de 1 a 5, a cada uma das expressões abaixo, avaliando a sua eficácia em transmitir o significado que pretendem comunicar.

Depois escolha uma palavra que você acredita que capture o significado da expressão.

1. Alguns carros são abacaxis. *

1 2 3 4 5

● ● ● ● ●

Agora escolha uma palavra que capture o significado da expressão acima: *

COMPLICAÇÃO

FACILIDADE

FRUTO

3.2 Results of norming studies

Results of both norming studies totaled 26,880 evaluations for the 84 metaphorical expressions constructed for this purpose: 13,608 evaluations performed by the 81 participants in Experiment 1, regarding the metaphors’ “Familiarity” and “Conventionality”; and 13,272, evaluations performed by the group of 79 participants in Experiment 2, who ranked them regarding “Aptness” and “Conventionality” – in both cases, these were performed in the subsequent research stage in order to construct the set of stimuli in the self-monitored reading experiment. Selection of higher rating metaphors concerning their “Conventionality” was based on the average of estimates obtained in both experiments.

Considering the individual, familiarity should be regarded as a discrete variable, involving exclusively two factors or “feelings”: the feeling of familiarity with “something”, and the feeling of “unfamiliarity” with such “something”, if what is perceived by the individual seems, respectively, familiar or not familiar to him/her. However, considering groups of individuals, it is permissible to admit degrees of familiarity, such as the “Familiarity” ratings obtained among participants in Experiment 1, which resulted in “unfamiliar”, “little familiar”, “moderately familiar”, “familiar” and “very familiar”, listed in Table 1.

Table 1 – Familiarity ratings in Experiment

Expression ratings	Answers YES (%) ⁸
Very familiar metaphors	between 80 and 100
Familiar metaphors	equal to or higher than 60 and lower than 80
Moderately familiar metaphors	equal to or higher than 40 and lower than 60
Little familiar metaphors	equal to or higher than 20 and lower than 40
Unfamiliar metaphors	lower than 20

Four expressions, that is to say, less than 5% of the total, were considered familiar by less than half of the participants: “Some neighbors are toads” (familiar to 41.98%), “*Alguns carros são abacaxis*” (“Some cars are lemons”) (40.74%), “*Algumas explicações são véus*” (“Some explanations are veils”) (39.51%), and the least familiar to the group, “*Alguns indivíduos são avestruzes*” (“Some individuals are ostriches”) (37.04%). The vast majority, 75 out of the 84 expressions (89% of the total), was considered familiar by 60% of the participants or more: from those that sounded familiar to a little bit more than half of the group, for instance, “*Alguns pensadores são águias*” (“Some thinkers are eagles”) (60.49%) and “*Algumas amigas são camaleões*” (“Some friends are chameleons”) (64.02%) to those that sound familiar to everyone, e.g.: “*Algumas sogras são cobras*” (“Some mothers-in-law are snakes”) (100%) and “*Algumas mulheres são furacões*” (“Some women are hurricanes”) (100%). This majority included metaphors ranked as “familiar” and “very familiar”, a pre-requisite for selecting stimuli for the self-monitored reading experiment to be performed in the subsequent research stage.

However, we decided to increase the minimum score to 85% of positive answers (YES) in judging the familiarity of expressions, maintaining 59 (or 70%) of the 84 candidates within the selection to make up the set of stimuli for the planned self-monitored reading experiment. Therefore, the selected metaphorical expressions were considered familiar by 85% of the participants, for instance, from “*Alguns políticos são raposas*” (“Some politicians are rats” – adapted to convey the idea of slyness) (85.19%), to “*Alguns professores são carrascos*” (“Some teachers are executioners” - adapted to convey the idea of mercilessness)

⁸ “Do the expressions below sound familiar to you?” Have you ever read or heard them?” “Check YES (if it is familiar) or NO (if it is not familiar)”. (see excerpt of “Familiarity” form in Figure 1).

(98.77%), to those that were unanimous in the group, e.g.: “Some drivers are snails” (100%) and “*Alguns músicos são feras*” (“Some musicians are animals” – adapted to convey the idea of greatness) (100%).

Table 2 – Aptness ratings in Experiment 2

Expression ratings	Scores (1 to 5) ⁹
Very high-apt metaphors	5
High-apt metaphors	4
Moderate-apt metaphors	3
Low-apt metaphors	2
Very low-apt metaphors	1

The Aptness ranking, in accordance with the scores between 1 and 5 given by the participants in Experiment 2 regarding the effectiveness of expressions, resulted in the following ratings: very low-apt, low-apt, moderate-apt, high-apt, and very high-apt, listed in Table 2.

From participants in Experiment 2, none of the 84 expressions received a score of lower than 2 for Aptness, which meant there were no low-apt or very low-apt ratings for the metaphors. Only 6, that is 7% of the total, received scores lower than 4, characterized as moderate-apt, namely: “*Alguns carros são abacaxis*” (“Some cars are lemons” – adapted to convey the idea of “problematic”) (with an average of 3.94), “*Alguns estudantes são traças*” (“Some students are bookworms”) (with 3.89), “*Algumas explicações são véus*” (“Some explanations are veils”) (3.86), “*Alguns vizinhos são sapos*” (Some neighbors are toads”) (3.77) and the two expressions with the lowest ratings, “*Alguns humanos são canários*” (“Some humans are canaries”) (3.65) and “*Alguns indivíduos são avestruzes*” (Some individuals are ostriches”) (3.57). In terms of “effectiveness”, almost all expressions – 78, or approximately 93% -- received scores between 4 and 5, limiting most expressions to the highest ratings, high-apt or very high-apt, e.g.: “*Alguns viciados são zumbis*” (“Some addicts are zombies”) (averaging 4.08), “*Alguns advogados são tubarões*” (“Some lawyers are sharks”) (4.32), “*Algumas estradas*

⁹ “Rate each of the expressions below from 1 to 5, evaluating their effectiveness in conveying the meaning they are intended to convey” (cf. excerpt of “Aptness” form, Figure 2).

são serpentes” (“Some roads are serpents”) (4.35), “*Alguns cunhados são malas*” (“Some brothers-in-law are a pain” – adapted to convey the idea of nuisance) (4.72), “*Algumas mães são leoas*” (“Some mothers are lionesses”) (4.82), and the one that was the closest to reaching a perfect score “*Algumas meninas são gatas*” (“Some girls are foxes” – adapted to convey the idea of attractiveness) (4.91).

Despite the high scores assigned to most expressions, we applied the same strict standards previously adopted, raising the minimum score of the metaphors’ aptness ratings to 4.5, thus preserving 61 out of the 84 (or 73%) of the candidates for stimuli in the planned self-monitored reading experiment. Therefore, the selected metaphorical expressions presented ratings of equal to or greater than 4.5, e.g.: “*Alguns policiais são gorilas*” (“Some police officers are gorillas”) (4.54), “*Alguns pais são bananas*” (“Some fathers are doormats” – adapted to convey the idea of submission) (4.67), as well as “*Alguns trabalhos são prisões*” (“Some jobs are jails”) (4.71), “*Algumas modelos são aviões*” (meaning “Some models are gorgeous”) (4.8), and those that were almost unanimous within the group “*Alguns lutadores são touros*” (“Some fighters are bulls”) (4.89) and “*Alguns mestres são espelhos*” (“Some masters are mirrors”) (4.89).

Average “Conventionality” estimates, obtained from the results of Experiments 1 and 2, translated into the following ratings: “new metaphors”, “little conventionalized metaphors”, “moderately conventionalized metaphors”, “conventionalized metaphors”, and “highly conventionalized metaphors” (Table 3). Estimates regarding “Conventionality”, collected in Experiments 1 and 2, were gathered in average estimates, weighted in relation to the number of participants in each of them in order to allow it to benefit from the totality of observations. The adoption of the same minimum level (85%) used to select expressions based on familiarity, in this study, limited the stimuli for the planned self-monitored reading experiment for highly conventionalized expressions.

Table 3 – “Conventionality” ratings in Experiments 1 and 2

Expression ratings	Congruent answers (%) ¹⁰
Highly conventionalized metaphors	between 80 and 100
Conventionalized metaphors	equal to or higher than 60 and lower than 80
Moderately conventionalized metaphors	equal to or higher than 40 and lower than 60
Little conventionalized metaphors	equal to or higher than 20 and lower than 40
New metaphors	lower than 20

Therefore, 43 expressions from the original set were selected to construct the stimuli for the self-monitored reading experiment, performed in the subsequent research stage. This selection included: “very familiar” metaphors (in at least 85% of the evaluations by participants in Experiment 1), very high-apt metaphors (with ratings equal to or higher than 4.5, obtained in the ranking in Experiment 2), and “highly conventionalized” metaphors (in 85% or more of the average estimates in the evaluations by the participants in Experiments 1 and 2) (see Table 4).

Table 4 – The 43 expressions selected in Experiments 1 and 2

The 43 expressions selected	English versions (* adapted)	Familiarity	Aptness	Conventionality
Alguns políticos são raposas.	Some politicians are rats.*	85.19%	4.67	97.45%
Algumas crianças são anjos.	Some children are angels.	98.77%	4.77	89.83%
Alguns chefes são toupeiras.	Some bosses are stupid.*	92.59%	4.66	94.92%
Alguns seguranças são armários.	Some security guards are monsters.*	93.83%	4.76	95.56%
Algumas modelos são aviões.	Some models are gorgeous.*	97.53%	4.8	98.73%
Alguns pais são bananas.	Some parents are doormats.*	92.59%	4.67	97.45%
Alguns maridos são cavalos.	Some husbands are jackasses.*	96.30%	4.68	100.00%
Algumas sogras são cobras.	Some mothers-in-law are snakes.	100.00%	4.82	99.36%
Alguns meninos são diabos.	Some boys are devils.	98.77%	4.67	96.84%
Algumas modas são febres.	Some fads are fevers.	96.30%	4.72	96.19%
Algumas senhoras são flores.	Some ladies are flowers.	91.36%	4.68	99.37%
Algumas mulheres são furacões.	Some women are hurricanes.	100.00%	4.81	98.11%
Algumas meninas são gatas.	Some girls are foxes.*	100.00%	4.91	100.00%

¹⁰ “[...] choose a word that captures the expression meaning” (see excerpts from the “Familiarity” and “Aptness” forms in Figures 1 and 2, respectively).

Algumas adolescentes são girafas.	Some teenagers are giraffes.	92.59%	4.63	99.36%
Alguns cordeiros são lobos.	Some sheep are wolves.	91.36%	4.68	96.83%
Algumas mães são leoas.	Some mothers are lionesses.	97.53%	4.82	98.11%
Alguns motoristas são lesmas.	Some drivers are snails.	100.00%	4.85	98.73%
Alguns genros são sanguessugas.	Some sons-in-law are leeches.	91.36%	4.66	98.73%
Alguns avós são santos.	Some grandparents are saints.	86.42%	4.65	87.97%
Algumas tias são vitrolas.	Some aunts are record players.	90.12%	4.72	98.09%
Algumas avós são corujas.	Some grandmothers are proud.*	98.77%	4.73	98.11%
Alguns cunhados são malas.	Some brothers-in-law are pains.*	97.53%	4.72	96.83%
Alguns músicos são feras.	Some musicians are animals.*	100.00%	4.84	100.00%
Algumas motos são foguetes.	Some motorcycles are rockets.	92.59%	4.85	99.37%
Alguns cozinheiros são porcos.	Some cooks are pigs.	97.53%	4.85	94.33%
Alguns locutores são papagaios.	Some announcers are parrots.	92.59%	4.72	96.20%
Algumas atitudes são máscaras.	Some attitudes are masks.	85.19%	4.61	92.42%
Algumas provas são torturas.	Some tests are torture.	98.77%	4.81	100.00%
Alguns botequins são chiqueiros.	Some bars are pigsties.	95.06%	4.85	96.81%
Alguns professores são carrascos.	Some teachers are executioners.	98.77%	4.87	96.20%
Algumas celebridades são antas.	Some celebrities are asses.	88.89%	4.75	99.36%
Alguns fumantes são chaminés.	Some smokers are chimneys.	96.30%	4.86	97.47%
Alguns barracos são fornos.	Some huts are furnaces.	93.83%	4.66	98.72%
Alguns caminhoneiros são tartarugas.	Some truck drivers are tortoises.	95.06%	4.85	98.72%
Algumas casas são fornalhas.	Some houses are stoves.	95.06%	4.72	97.47%
Algumas garotas são violões.	Some girls are curvaceous.*	95.06%	4.8	97.44%
Alguns amigos são pilares.	Some friends are pillars.	87.65%	4.77	96.84%
Algumas torcidas são quadrilhas.	Some team fans are gangs.	92.59%	4.63	89.89%
Algumas críticas são coices.	Some criticisms are kicks.	92.59%	4.72	94.30%
Alguns lutadores são touros.	Some fighters are bulls.	93.83%	4.89	98.08%
Alguns mestres são espelhos.	Some masters are mirrors.	93.83%	4.89	98.73%
Alguns trabalhadores são formigas.	Some workers are ants.	92.59%	4.62	86.08%
Algumas vizinhas são moscas.	Some neighbors are flies.	93.83%	4.51	95.53%

4 Experiment 3: self-monitored reading of BP metaphors

The self-monitored reading experiment (self-paced, non-cumulative, moving-window reading) was designed to evidence direct processing supported by the Class-inclusion theoretical model (GLUCKSBERG; KEYSAR, 1990), for very familiar, very high-apt, and highly conventionalized metaphors in BP, selected in norming studies conducted in our research, observing the reading times impact upon the “Recognition Problem” locus of the metaphorical meaning, that is, the vehicle for each of the expressions (e.g.: “*um furacão*” (“a hurricane”) in “*Irene é um furacão*” (“Irene is a hurricane”) demanded metaphorical interpretation). The reading times for the corresponding fragments were also recorded, e.g.: “*um furacão*” (“a hurricane”) in “*Irene é um furacão*” was to be interpreted literally, as was the reading time for the control fragments, e.g.: “*uma jovem*” (“a young woman”) in literal class inclusion statements, in this case, “*Irene é uma jovem*” (“Irene is a young woman”).

Participants: 66 volunteers participated in the experiment: 20 men and 46 women, an average of 27 years of age. Thirty-five volunteers were recruited among undergraduate students from the Languages Department at Federal Fluminense University (UFF) and 3 were administrative workers at the institution, the latter with college degrees. Twenty-eight volunteers were students in different areas at the Severino Sombra University (Nursing, Production Engineering, Chemical Engineering, Medicine, and Psychology). Among these, 24 were undergraduate students and 4 were specialists.

Materials: 9 triplets of paragraphs (transcribed in the Appendix), equally divided into 11 fragments/syntagmas, were prepared to include: in type 1: expressions type “X is a Y” constructed based on metaphors selected in the pre-tests, e.g.: “*Irene é um furacão*” (“Irene is a hurricane”), constructed based on “*Algumas mulheres são furacões*” (“Some women are hurricanes”) (see Table 4); in type 2: expressions with literal meaning of the same type (X is a Y), e.g.: “*Irene é um furacão*”, constructed based on what had been learned from hurricane Irene, which struck the Caribbean and the USA in 2011; and in type 3: class inclusion statements of the same type (X is a Y), e.g.: “*Irene é uma jovem*” (“Irene is a young woman”), categorizing the topic of the expression. In type 1 contexts, the expressions “X is a Y” required metaphorical interpretations; in type 2 contexts, the same expressions (X is a Y) required literal interpretations, and in type 3 contexts, “X is

a Y” were usual literal class inclusion statements. Critical fragments of the - “Y” in all - metaphorical and literal expressions and of literal inclusion statements, were balanced as the number of syllables, 4 (7th segment); and in type and number of constituents, “Det + N” (in the examples, respectively, “*um furacão*”, “*um furacão*”, “*uma jovem*”). In addition, all the segments of the 9 triplets were equally balanced in number of syllables before the critical fragment, and up to the ninth one. After reading each paragraph, participants answered a question about the content read, by pressing the buttons corresponding to YES or NO on the button box attached to the stimuli exhibition device. The bars in the examples below, on paper, limit the 11 fragments provided in each of the reading windows, at the rhythm imposed by each of the participants upon reading the paragraphs.

Type 1 context: requiring metaphorical interpretation

Com muita frequência, / mulheres brasileiras / trabalham fora, / têm filhos / e estudam. / Irene é / um furacão / que empolga / todo mundo / no trabalho, em casa / e na faculdade.

(“Quite frequently, / Brazilian women / have jobs, / have children / and study. / Irene is / a hurricane / who excites / everyone / at work, at home, / and at school.”)

Type 2 context: requiring literal interpretation

Com muita frequência, / fenômenos naturais / recebem nomes / de gente / dos mais comuns. / Irene é / um furacão / que assolou / o Caribe / e o leste dos EUA / em 2011.

(“Quite frequently, / natural phenomena / are named / after / common people. / Irene is / a hurricane / that struck / the Caribbean / and the eastern USA / in 2011.”)

Type 3 context: requiring interpretation of the class inclusion statement

Com muita frequência, / mulheres brasileiras / trabalham fora, / têm filhos / e estudam. / Irene é / uma jovem / que empolga / todo mundo / no trabalho, em casa / e na faculdade.

("Quite frequently, / Brazilian women / have jobs, / have children, / and go to school. / Irene is / a young woman / who excites / everyone / at work, at home, / and at school.")

Stimuli were distributed (within subjects, counterbalanced in Latin squares) so that participants read paragraphs in all 3 conditions (by reading 3 stimuli per condition and never reading the same stimulus for more than one condition). Eighteen additional distractive paragraphs were interspersed among the 9 target-paragraphs, and the presentation order for the total 27 paragraphs was randomized in each session with each participant. Questions after each of the target-paragraphs required participants to provide approximately as many YES answers as NO answers.

Procedures: In individual sessions, each participant read the stimuli, fragment by fragment, prompted, at the participant's command, on an Apple computer screen, by pressing the "next screen button" from a button box attached to the device. For this, the stimuli presentation followed a protocol built in PsyScope, so as to record reading times and participants' answers. Prior to each session, participants underwent training, consisting of reading 4 stimuli, in order to familiarize themselves with equipment operation and with the experiment routine. Undergraduate students were given a certificate for participating in extra-curricular activities, complementary to the studies they might have been developing. At the end of each session, participants were told the experiment purposes, and any potential doubts they may have had were clarified.

Hypothesis: According to the theoretical class-inclusion model (GLUCKSBERG; KEYSAR, 1990), the very familiar, highly conventionalized, and very high-apt BP metaphors selected during the norming studies should be processed directly, without mediation by the derivation from literal meaning, as are literal class inclusion statements. Alternatively, according to the so-called Standard Pragmatic Model, the same metaphors should be processed indirectly, based on the derivation of the expression's literal meaning and its rejection in the utterance context.

Predictions: Due to the research hypothesis, of direct processing of very familiar, highly conventionalized, and high-apt metaphors, significant differences are not expected between the reading time for these

metaphor vehicles and for the corresponding nominal syntagmas, both in the equivalent literal meaning expressions and in the class inclusion literal statements. If indirect processing prevails, significantly higher average reading times will most likely be observed for the metaphor vehicles than those for the same target-nominal syntagmas in literal interpretation and the target-nominal syntagmas in class inclusion literal statements.

4.1 Self-monitored reading experiment result analysis

We hereby adopt the methodological strictness that, according to Janus and Bever's criticism (1985), is missing in the literature to date, in supporting as well as in rejecting indirect processing of the metaphor that follows the Standard Pragmatic Model. For such, we divided the stimuli into syntagmas and measured the reading times (RT) on the "Recognition Problem" locus of the metaphorical meaning, the vehicles of metaphors. However, unlikely in Janus and Bever (1985), whose findings suggest that new metaphors are processed indirectly, based on the derivation of the expression's literal meaning, we obtained results that support the hypothesis of direct processing of conventionalized, familiar, and high-apt metaphors. In fact, Janus and Bever (1985, p. 485) admit the possibility of direct processing of what they call frozen metaphors, and our findings for BP are apparently no different from that.

Contrary to the predictions associated with the hypothesis of indirect processing, in terms of the Standard Pragmatic Model, participants in the self-monitored reading experiment conducted by us did not spend significantly longer periods in reading the metaphor vehicles, e.g., "*um furacão*" ("a hurricane"), in "*Irene é um furacão*" ("Irene is a hurricane"), than they spent reading the same nominal syntagmas in expressions with literal interpretation, e.g., "*Irene é um furacão*" ("Irene is a hurricane"), referring to the hurricane that struck the Caribbean in 2011. They did not spend longer periods reading the nominal syntagmas corresponding to the literal class inclusion statements, e.g., "*Irene é uma jovem*" ("Irene is a young woman" - adapted), thus suggesting the validity of the hypothesis of direct processing via Glucksberg and Keysar's (1990) Class inclusion/Dual reference, as demonstrated by the average RT listed in Table 5.

Table 5 – Average reading times (RT in ms) for critical fragments (highlighted) under the following conditions: Literal, Metaphorical and Literal Class Inclusion Statement (LCIS)

Contexts	RT	Examples (paragraph extracts)
Literal	954.25	<i>[...] fenômenos naturais recebem nomes de gente. Irene é um furacão que assolou o Caribe [...].</i> (“[...] natural phenomena are named after people. Irene is a hurricane that struck the Caribbean [...].”)
Metaphorical	925.84	<i>[...] mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é um furacão que empolga todo mundo [...].</i> (“[...] Brazilian women have job, have children, and go to school. Irene is a hurricane who excites everyone [...].”)
LCIS	867.89	<i>[...] mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é uma jovem que empolga todo mundo [...].</i> (“[...] Brazilian women have job, have children, and go to school. Irene is a young woman who excites everyone [...].”)

The RT for critical fragments did not exhibit normal distribution for any of the 3 conditions: Literal, Metaphor, or Literal Class Inclusion Statement (LCIS) (for Literal condition, Kolmogorov-Smirnov statistics=0.14, $p < .05$; for Metaphor condition, Kolmogorov-Smirnov=0.12, $p < .05$; and for the LCIS condition, Kolmogorov-Smirnov=0.12, $p < .05$). In addition, significant differences were not detected in Multiple Comparisons among the conditions, LCIS-Literal, LCIS-metaphor, Literal-Metaphor, in statistical analysis per subject (Kruskal-Wallis $X^2(2)=2.95$, $p > .05$) and in the analysis per item (Kruskal-Wallis $X^2(2)=3.55$, $p > .05$). The Test Power used was estimated at 70%, capable of evaluating differences of, at least, 200 milliseconds (ms) between the RTs of the critical fragments for the conditions and differences, such as those Janus and Bever (1985) reported and that were not present in this case.¹¹

¹¹ The epistemological nature of findings such as ours, which are based on the absence of statistically significant differences among experimental conditions, has been questioned, traditionally, under the terms of the so-called argumentum *ad ignorantiam* (the absence of evidence is not evidence of the absence). Evidence of absence, in the tradition that goes back to Sextus Empiricus’ (2nd century) skepticism, should be considered fallacious, as absence of evidence that may simply have not been found.

Participants' answers to the questions on the content in the paragraphs satisfactorily attested to the involvement of individuals in performing the proposed task. Considering the totality of participants (66), we detected the vast majority (94%) of correct answers. What's more, approximately 99% of the participants answered 90% of the questions correctly, ensuring that the results are not attributed to the lack of attention by readers or to their difficulty in understanding stimuli.

The statistically non-significant difference (only 28.4 ms) between the average reading time (average RT) for metaphor vehicles and for the nominal syntagmas interpreted literally suggests that familiar, high-apt, and conventionalized metaphorical meanings (thus considered in the participants' off-line answers in the norming studies) are indeed "available" when the online structuring of predicative expressions and their interpretation in appropriate contexts take place. In relation to this difference (statistically non-significant, of 28 ms), it may be asked whether they caused any "strangeness" and difficulty in interpreting literal expressions that, in the stimuli used, assigned "people's names" to non-human beings, whether animate or inanimate, e.g.. "*Irene é um furacão*" ("Irene is a hurricane") in the paragraph "[...] *fenômenos naturais recebem nomes de gente. Irene é um furacão que assolou o*

However, there has been an increase in the view that some cases in which a proposition is assumed to be true simply because it may not be proven false, or is false because it may not be proven true, are not considered fallacies (WALTON, 1992, p. 381-4). Among such cases, circumstances are admitted in which conclusive evidence does not appear, despite being sought; and the proposition that expresses the absence of such results is assumed to be true or false, although temporarily, considering the possibility of future rejections (p. 383). Therefore, we argue that this is the case of differences between the RT of critical segments for the conditions (as an Alternative Hypothesis) that were not evidenced in the experiment we performed, enticing the defense - at least presumptive defense, under Walton's terms - of the Research Hypothesis. To add further support to the assumption that the experimental findings we obtained are not fallacious, we adduce that, if, on the one hand, the absence of significant differences among RT for critical segments under the 3 conditions may be paraphrased under the terms of non-rejection of the null hypothesis, on the other hand, the test power used in the statistic treatment of data, estimated at 70%, would be capable of surprise differences of, at least, 200 ms, such as those reported by Janus and Bever (1985) in the processing of new metaphors, and, therefore, respond to the challenge that a type II error may be occurring here, that is, not rejecting the false H₀. (COOLICAN, 2014, p. 427-8).

Caribe [...]” (“[...] natural phenomena are named after people. Irene is a hurricane that struck the Caribbean [...])” (see Appendix for the other references). This average RT for the literal interpretation of nominal syntagmas also departed (86.36 ms, a difference that was also not statistically significant) from the average RT of nominal syntagmas in the class inclusion statements, e.g., “*Irene é uma jovem*” (“Irene is a young woman”), which are part of BP speakers’ repertoire, as consolidated semantic knowledge. If, therefore, some literal expressions cause “strangeness” (and literal expressions should not cause it), it should be argued that the effect obtained by confronting, e.g., “*Fernanda é uma lesma*” (“Fernanda is a snail”), in the paragraph “*Em certas culturas, animais exóticos são adotados e ganham apelidos. Fernanda é uma lesma que habita um viveiro na casa de criadores nas Filipinas*” (“In some cultures, exotic animals are adopted and given nicknames. Fernanda is a snail that lives in a nursery in a breeder’s in the Philippines”), is due, in such cases, to the pragmatic “misfortune” of the literal statement, rather than a breach in the sentence’s conditions of truth (see also Matthewson, 2004, p. 409).¹² It is thus true that problems, in principle, inherent in “performative” utterances would be extended to the examples of the so-called “constative” utterances, although this would be endorsed by the literature. At least for Gerken (1994, p. 78), by “describing”, “explaining”, and “theorizing” we also commit to performative activities, which means stating that it would be more appropriate to consider “constative” or “descriptive” as special cases of “performative.” In Petrey (1990), it is possible to determine that even Austin (1962) recognized that the distinction (originally proposed by Austin) was too strong and leaned towards reevaluating it under the terms explained above. In Grandesso (2006), we are able to get in contact with the literature in the Portuguese language that addresses the views exposed herein.

Regarding the hypothesis of the direct processing of metaphors, we return to the key issue to be highlighted: the average RT for metaphor vehicles did not significantly differ from the average RT for the same nominal syntagmas interpreted literally. This did not significantly differ

¹² We thank Dr. Luciana Sanchez Mendes, our colleague at GEPEX – UFF, for the suggestion to direct, towards the path above, the evaluation of the instigating (albeit statistically marginal) difference between the reading times of nominal syntagmas for literal interpretation and for literal class inclusion statements.

from the average RT of the corresponding nominal syntagmas in the literal class inclusion statements (57.95 ms), which leads to the defense of Glucksberg and Keysar's (1990) position that metaphors are what they appear to be: class inclusion statements, metaphorical categories that are the names of the vehicles, which are prototypical members of such categories. The fact that, in the BP data reported here, no significant differences between the RT of critical fragments for the conditions suggests that familiar, high-apt, and conventionalized metaphors are processed as class inclusion statements, according to the Class-inclusion model (GLUCKSBERG; KEYSAR, 1990).

The absence of significant differences between the RT obtained could result, perhaps, from the sample composition. Among the 66 participants in the experiment, there was no actual balance in the distribution between men (1/3) and women (2/3); the individuals selected were originally from different cities (although two were from the same state); and were students or had graduated from two different higher education institutes, a public university, and a private college; they were undergraduate students (59), had undergraduate degrees (1), and had specialist degrees (6); in different majors.

We certainly did not mean to ignore the fact that the understanding of metaphors by individuals within different sociocultural contexts may differ considerably (for an introduction to the studies on the cultural aspects involved in understanding metaphors, refer to Gibbs Jr. (2008) and Ortony (1993 [1979]) and the references recommended in it). So much so that, during the first research stage, evaluations were collected about familiarity, aptness, and conventionality of expressions that originated the stimuli used in the self-monitored reading experiment.

Nevertheless, from that point on, we started to systematize findings from the self-monitored reading experiment under the light of what may be called the "standardization" of the global sample. By restricting the sample to the participants who provided right answers to all questions (100% instead of 94%); and to those who attended UFF and were in the second semester majoring in Languages – we reduced the total to 38 individuals whose average age (previously, 27.3 years, and now, 26.6 years) was altered by less than 1 year; and to a subset whose ratio between men (25%) and women (75%) reflects the public majoring in Languages in the country. If we consider now, the "standardized" sample, there are no reasons to question the results previously shown (see Table 6), which did not

diverge from those previously exhibited (in Table 5, when the total of 66 participants was considered), and, therefore, do not justify other conclusions.

Table 6 – Average reading times (RT) for critical fragments under the conditions: Literal, Metaphor, Literal Class Inclusion Statement (in ms), with 66 and with 38 participants

	Literal	Metaphorical	Control
66 participants	954.25	925.84	867.89
38 participants	1008.02	968.11	930.11

Consequently, one way or another, the hypothesis of the indirect processing of familiar, high-apt, and conventionalized BP metaphors is not supported by the results of the self-monitored reading experiment conducted in this research. Average RT did not differ significantly, thus suggesting that participants, both in the 66 of the original sample and in the 38 in the subset, belonging to a geographically and socially more homogeneous group, understood, with no additional difficulties in relation to the same literal expressions and literal class inclusion statements, the metaphors inserted in appropriate contexts. By interpreting the metaphorical meaning of, e.g., “*Irene é um furacão*” (“Irene is a hurricane”) (because apparently Irene disrupts the social environment wherever she goes), native speakers of BP in the selected sample did not find problems, in the same way that they did not find trouble with the literal interpretation of “*Irene é um furacão*” (“Irene is a hurricane”) (the hurricane that struck the Caribbean). They had no trouble interpreting the metaphorical meaning of the same expressions that were greater than the trouble they had to interpret “*Irene é uma jovem*” (“Irene is a young woman”), literal class inclusion statement (including Irene in the category of ‘*jovem*’ (‘young woman’), which enables us to infer, as proposed by Glucksberg et al. (GLUCKSBERG; KEYSAR, 1990; GLUCKSBERG, 1998; GLUCKSBERG, 2003), that they understood the metaphors as well as they understood the literal class inclusion statements.

Figure 3 – Properties associated with the metaphorical meaning M-furacão of “*furacão*” (“hurricane”) at the supraordinate level, and to the literal meaning L-furacão “*furacão*” (“hurricane”) at the basic category level

Nível Supraordenado da Categoria:	
propriedades M- furacão:	Muito impetuoso(a) Avassalador(a) Muito agitado(a) Muito Forte
Nível básico da Categoria:	
propriedades L- furacão:	Muito impetuoso(a) Avassalador(a) Muito agitado(a) Muito Forte Destruidor(a) Muito veloz Devastador(a) Tumultuador(a)

As Glucksberg (1998, p. 41) hierarchically structures, in 2 reference levels, a basic level and a supraordinate or metaphorical level, to account for “my lawyer is a shark”, the metaphorical category “*furacão*” (“hurricane”) is illustrated in Figure 3.

According to the Class-inclusion/Dual reference model, before the expression “*Irene é um furacão*” (“Irene is a hurricane”), the listener/reader promptly accesses the metaphorical meaning of the metaphor vehicle, “*furacão*” (“hurricane”) – conventionalized (in BP, even included in the dictionary). In the Glucksberg and Keysar’s model (1990) of direct metaphor processing by categorization, metaphors such as “*Irene é um furacão*” (Irene is a hurricane”) would be understood as literal class inclusion statements, due to the inclusion of topics from pre-existing categories created from conventionalized vehicles.

5 Conclusions

The results of the experiments performed support the hypothesis defended by Glucksberg and Keysar (1990), of direct processing of familiar, high-apt metaphors, whose vehicles are conventionalized,

according to the Class-inclusion model. RTs were measured for the critical fragments for the expressions fit in appropriate contexts, and no significant difference was identified between the RT of metaphor vehicles and the equivalent literal expressions. A significant difference was also not identified between the average RT of metaphor vehicles and literal class inclusion statements, which suggests that the metaphors in question were processed in accordance with the Class-inclusion/Dual reference (GLUCKSBERG; KEYSAR, 1990).

It is important to highlight here that we did not consider extending the idea of direct processing via Class-inclusion to categories other than the pre-existing ones, as aimed by Glucksberg (2008). For Glucksberg (2008), the conventionality of “vehicles” results necessarily from their use in well-constructed expressions. Therefore, high-apt metaphors, despite being new, would be interpreted as class inclusion statements. In the future, such a hypothesis may (and should) be duly tested.

Despite its originality, the present study does not cross the original limits defined for it: that of evaluating in BP the processing of conventionalized, familiar, and high-apt metaphors. Still awaiting the interest of future research, the study of BP metaphor processing with the assistance of eye movement tracking technology in understanding metaphors, as well as the extraction of the potential evoked during the hearing/reading of BP metaphors and investigations based on neuroimaging.

Acknowledgments

We wish to thank the anonymous peers for identifying aspects in the text that deserved to be reformulated. We thank Reviewer 1 for the care towards the transition to objectives, hypothesis, and predictions of the self-monitored reading experiment. We thank Reviewer 2 for advocating (but not demanding) the use of Portuguese equivalents for the English terms that are established in the international literature, which will happen more often, as the field of metaphor processing becomes more consolidated among us. Finally, we thank the volume organizer for urging us to hold the necessary discussion (see note 12) on the findings we reported, considering the ideas surrounding the so-called *argumentum ad ignorantiam*.

References

- AULETE, Caldas. *Aulete Digital*: o dicionário da língua portuguesa na internet. Versão online. Disponível em: <<http://www.aulete.com.br/>>. Acesso em: 16 set. 2015.
- AUSTIN, J.L. *How to do things with words*. Cambridge, Mass.: Harvard University Press, 1962.
- BLASKO, D.; CONNINE, C.M. Effects of familiarity and aptness on metaphor processing. *The Journal of Experimental Psychology: Learning, Memory, and Cognition*, Washington, v. 19, p. 295-308, 1993. <https://doi.org/10.1037/0278-7393.19.2.295>
- BOWDLE, B. F.; GENTNER, D. The career of metaphor. *Psychological Review*, American Psychological Association, v. 112, n. 1, p. 193-216, 2005. <https://doi.org/10.1037/0033-295x.112.1.193>
- CHIAPPE, D.; KENNEDY, J. M.; SMYKOWSKI, T. Reversibility, aptness, and the conventionality of metaphors and similes. *Metaphor and Symbol*, v. 18, n. 2, p. 85-105, 2003. https://doi.org/10.1207/S15327868MS1802_2
- CLARK, H. H.; LUCY, P. Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, Elsevier B. V., n. 14, p. 56-72, 1975.
- COOLICAN, H. *Research methods and statistics in psychology*. 6th ed. London; New York: Psychology Press; Taylor & Francis Group, 2014.
- CRAIG, Colette G. (Ed.). *Noun classes and categorization*: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983. John Benjamins Publishing Company, 1986.
- DENNY, J. Peter. The semantic role of noun classifiers. *Noun classes and categorization*, p. 297-308, 1986.
- DULCINATI, G.; MAZZARELLA, D.; POUSSCOULOUS, N.; RODD, J. Processing metaphor: The role of conventionality, familiarity and dominance. *UCL Working Papers in Linguistics*, London, University College London, v. 26, p. 72-88. 2014.
- FERREIRA, Aurélio Buarque de Hollanda. *Miniaurélio*: o Minidicionário da Língua Portuguesa. 6. ed. ver. ampl. Curitiba: Positivo, 2004.
- FERREIRA, Aurélio Buarque de Hollanda. *Novo Aurélio século XXI*: o dicionário da língua portuguesa. 3. ed. ver. ampl. Rio de Janeiro: Nova Fronteira, 1999.

FRASER, B. The interpretation of novel metaphors. In: ORTONY, A. (Ed.). *Metaphor and thought*. Cambridge: Cambridge University Press, p. 329-341, 1993. <https://doi.org/10.1017/CBO9781139173865.017>

GARRET, J. *Aristotle on metaphor*. 2007. Disponível em: <<http://people.wku.edu/jan.garrett/401s07/arismeta.htm>>. Acesso em: 6 dez. 2016.

GERKEN, L. A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, Cambridge, Cambridge University Press, v. 21, p. 565-584, 1994. <https://doi.org/10.1017/s0305000900009466>

GIBBS JR, R. W. Contextual effects in understanding indirect requests. *Discourse Processes*, v. 2, p. 1-10, 1979. <https://doi.org/10.1080/01638537909544450>

GIBBS JR, R. W. *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, 2008. <https://doi.org/10.1017/CBO9780511816802>

GIBBS JR, R. W. Your wish is my command: Convention and context in interpreting indirect requests. *Journal of Verbal Learning and Verbal Behavior*, Elsevier B. V., v. 20, p. 431-444, 1981.

GLUCKSBERG, S. How metaphors create categories – quickly. In: GIBBS JR, R. W. *The Cambridge Handbook of Metaphor and Thought*. Cambridge: Cambridge University Press, p. 67-83, 2008. <https://doi.org/10.1017/CBO9780511816802.006>

GLUCKSBERG, S. The Psycholinguistics of metaphor. *Trends in Cognitive Sciences*, Elsevier, v. 7, n. 2, p. 92-96, 2003. [https://doi.org/10.1016/S1364-6613\(02\)00040-2](https://doi.org/10.1016/S1364-6613(02)00040-2)

GLUCKSBERG, S. Understanding metaphors. *Current Directions in Psychological Science*, Cambridge, Cambridge University Press, v. 7, n. 2, p. 39-43, 1998.

GLUCKSBERG, S.; GILDEA, P.; BOOKIN, H. B. On understanding nonliteral speech: can people ignore metaphors? *Journal of Verbal Learning and Verbal Behavior*, Elsevier B. V., v. 21, p. 85-98, 1982.

GLUCKSBERG, S.; KEYSAR, B. Understanding Metaphorical Comparisons: Beyond Similarity. *Psychological Review*, American Psychological Association, v. 97, n. 1, p. 3-18, 1990. <https://doi.org/10.1037/0033-295x.97.1.3>

- GRANDESSO, M. *Sobre a reconstrução do significado: uma análise epistemológica e hermenêutica da prática clínica*. 2. ed. São Paulo: Casa do Psicólogo, 2006.
- GRICE, H. P. Logic and Conversation. In: COLE, P.; MORGAN, J. L. (Ed.). *Syntax and Semantics*. New York: Academic Press, 1975. [v. 3: Speech Acts, p. 41-58].
- HAGE, P.; MILLER, W. R. 'Eagle' = 'bird': A note on the structure and evolution of Shoshoni ethnoornithological nomenclature. *American Ethnologist*, 3, p. 481-488, 1976. <https://doi.org/10.1525/ae.1976.3.3.02a00060>
- HARRIS, R. J. Comprehension of metaphors: A test of the two-stage processing model. *Bulletin of the Psychonomic Society*, v. 8, n. 4, p. 312-314, 1976. <https://doi.org/10.3758/BF03335150>
- HOUAISS, A.; VILLAR, M. S. *Dicionário Houaiss da Língua Portuguesa*. 3. ed. Rio de Janeiro: Objetiva, 2009.
- JANUS, R. A.; BEVER, T. G. Processing of metaphoric language: an investigation of the three-stage model of metaphor comprehension. *Journal of Psycholinguistic Research*, Springer International Publishing A. G., v. 14, n. 5, p. 473-487, 1985.
- JOHNSON, A.T. Comprehension of metaphors and similes: a reaction time study. *Metaphor and Symbolic Activity*, Lawrence Erlbaum Associates, Inc., v. 11, n. 2, p. 145-159, 1996.
- JONES, L. L.; ESTES, Z. Roosters, robins, and alarm clocks: Aptness and conventionality in metaphor comprehension. *Journal of Memory and Language*, Elsevier, v. 55, n. 1, p. 18-32, 2006. <https://doi.org/10.1016/j.jml.2006.02.004>
- MATTHEWSON, L. On the methodology of semantic fieldwork. *International Journal of American Linguistics*, Chicago, University of Chicago, v. 70, n. 4, p. 369-415, Oct. 2004.
- MILLER, G. A. Images and models: similes and metaphors. In: ORTONY, A. (Ed.). *Metaphor and thought*. Cambridge: Cambridge University Press, p. 202-250, 1979.
- MORISON, Benjamin. Sextus Empiricus. In: ZALTA, Edward N. (Ed.). *The Stanford Encyclopedia of Philosophy*. Stanford: Spring Edition, 2014. Disponível em: <<https://plato.stanford.edu/archives/spr2014/entries/sextus-empiricus/>>. Acesso em: 1 mar. 2017.

ORTONY, A. *Metaphor and thought*. 2nd edition. New York: Cambridge University Press, 1993 [1979]. <https://doi.org/10.1017/CBO9781139173865>

ORTONY, Andrew et al. Interpreting metaphors and idioms: Some effects of context on comprehension. *Journal of verbal learning and verbal behavior*, v. 17, n. 4, p. 465-477, 1978.

PETREY, S. *Speech Acts and Literary Theory*. New York; London: Routledge, 1990.

RIBEIRO, A. J. C.; RICCI, A.Q.; KENEDY, E. O processamento psicolinguístico da metáfora: um estudo inédito no português brasileiro. In: ENCONTRO NACIONAL DA ANPOLL (ENANPOLL), XXXI, 2016, Campinas. *Poster*. Campinas: UNICAMP, 2016.

RICCI, A. Q. *O processamento psicolinguístico da metáfora: um estudo experimental no PB*. 2016. 75f. Dissertação (Mestrado em Estudos de Linguagem) – Instituto de Letras, Universidade Federal Fluminense, 2016.

SEARLE, J. Metaphor. In: ORTONY, A. *Metaphor and thought*. 2nd edition. New York: Cambridge University Press, 1993. p. 83-111. [1979]

SHINOFF, P. Demjanjuk war-crimes tribunal strikes deep fear among Jews. *San Francisco Examiner*, June 14, p. A8, 1987.

STROOP, J.R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, American Philosophical Association, v. 18, p. 643-662, 1935.

THIBODEAU, P. H.; DURGIN, F. H. Metaphor aptness and conventionality: A processing fluency account. *Metaphor and Symbol*, Francis & Taylor Online, v. 26, p. 206-226, 2011.

TRAGER, G. L. “Cottonwood-Tree”, a south-western linguistic trait. *International Journal of American Linguistics*, 9, p. 117-118, 1938. <https://doi.org/10.1086/463823>

TVERSKY, A. Features of similarity. *Psychological Review*, American Psychological Association, v. 84, p. 327-352, 1977. <https://doi.org/10.1037/0033-295x.84.4.327>

WALTON, Douglas. Nonfallacious arguments from ignorance. *American Philosophical Quarterly*, v. 29, n. 4, Oct. 1992. Disponível em: <<http://dougwalton.ca/papers%20in%20pdf/92nonfallacious.pdf>>. Acesso em: 1 mar. 2017.

Appendix: Experimental paragraphs for the chronometric experiment

The experimental paragraphs below are organized in the sequence Literal/Metaphor/Literal Class Inclusion Statement, for each triplet. The translations of some expressions were adapted (*) to retain their original ideas.

1. *Com muita frequência, fenômenos naturais recebem nomes de gente dos mais comuns. Irene é um furacão que assolou o Caribe e o leste dos EUA em 2011.* (“Quite frequently, natural phenomena are given the most common people’s names. Irene is a hurricane that struck the Caribbean and the eastern US in 2011.”)

Com muita frequência, mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é um furacão que empolga todo mundo no trabalho, em casa e na faculdade. (“Quite frequently, Brazilian women have jobs, have children, and go to school. Irene is a hurricane who excites everyone at work, at home, and at school.”)

Com muita frequência, mulheres brasileiras trabalham fora, têm filhos e estudam. Irene é uma jovem que empolga todo mundo no trabalho, em casa e na faculdade. (“Quite frequently, Brazilian women have jobs, have children, and go to school. Irene is a young woman* who excites everyone at work, at home, and at school”)

2. *Excepcionalmente, móveis italianos têm qualidade e nome no mercado. Bérghamo é um armário que recebe dos clientes ótimas avaliações nos comentários.*

Excepcionalmente, segurança de rua porta armas para ser respeitado. Bérghamo é um armário que trabalha com ajuda do porte avantajado e da força bruta. (“Exceptionally, security guards carry guns to be respected. Bérghamo is a monster* who works with the help of his massive size and raw strength.”)

Excepcionalmente, segurança de rua porta armas para ser respeitado. Bérghamo é um vigia que trabalha com ajuda do porte avantajado e da força bruta. (“Exceptionally, security guards carry guns to be respected. Bérghamo is a security guard who works with the help of his massive size and raw strength.”)

3. *Em certas culturas, animais exóticos são adotados e ganham apelidos. Fernanda é uma lesma que habita um viveiro na casa de criadores nas Filipinas.* (“In certain cultures, exotic animals are adopted and are given nicknames. Fernanda is a snail that lives in a nursery in a breeder’s house in the Philippines”)

Costumeiramente, motoristas idosos guiam devagar sem razão no trânsito. Fernanda é uma lesma que atrasa a chegada no trabalho ou lazer quando usa carro. (“Older drivers usually drive slowly in traffic for no reason. Fernanda is a snail that is late to get to work or recreation when she uses a car.”)

Costumeiramente, motoristas idosos guiam devagar sem razão no trânsito. Fernanda é uma chofer que atrasa a chegada no trabalho ou lazer quando usa carro. (“Older drivers usually drive slowly in traffic for no reason. Fernanda is a driver that is late getting to work or recreation when she uses a car.”)

4. *Em certas famílias, bichos de estimação distanciam-se do padrão esperado. Tereza é uma cobra que só pica raramente e não tem qualquer veneno, segundo o dono.* (“In some families, pets are different from the expected standard. Tereza is a snake that rarely bites and is not poisonous, according to the owner.”)

Em certas famílias sogras problemáticas implicam muito com genros, sem motivos. Tereza é uma cobra que ataca sem aviso o marido da filha em qualquer lugar. (“In some families, problematic mothers-in-law pick on their sons-in-law for no reason. Tereza is a snake who attacks her daughter’s husband anywhere, without warning.”)

Em certas famílias sogras problemáticas implicam muito com genros, sem motivos. Tereza é uma sogra que ataca sem aviso o marido da filha em qualquer lugar. (“In some families, problematic mothers-in-law pick on their sons-in-law for no reason. Tereza is a mother-in-law who attacks her daughter’s husband anywhere, without warning.”)

5. *Em alguns lugares, primatas em extinção ganham proteção de ONGs e pessoas. Samuel é uma fera que resiste na África à extinção dos gorilas até o momento.* (“In some places, endangered primates are protected by NGOs and individuals. Samuel is a beast that has resisted the extinction of gorillas in Africa to date.”)

Por todo o mundo, guitarristas famosos sobem em palcos ou tocam ao ar livre. Samuel é uma fera que fascina o público da banda de rock Skank há muitos anos. (“All over the world, famous guitar players play on stages or in open environments. Samuel is an animal who has fascinated the audience of the rock band Skank for many years.”)*

Por todo o mundo, guitarristas famosos sobem em palcos ou tocam ao ar livre. Samuel é um músico que fascina o público da banda de rock Skank há muitos anos. (“All over the world, famous guitar players play on stages or in open environments. Samuel is a musician who has fascinated the audience of the rock band Skank for many years.”)

6. *Em filmes infantis, grupos de felinos formam famílias felizes e saudáveis. Maria é uma gata que alegra as crianças na famosa animação Os Aristogatas. (“In children’s movies, groups of felines are happy and healthy families. Marie is a cat that entertains children in the famous animation “The Aristocats.”)*

Em qualquer época, meninas sedutoras estão no centro de festas e baladas. Maria é uma gata que encanta os garotos em qualquer reunião de adolescentes. (“Seductive girls are the center of attention in parties and nightclubs at any time. Marie is a fox who charms boys in any teenage gathering.”)*

Em qualquer época, meninas sedutoras estão no centro de festas e baladas. Maria é uma moça que encanta os garotos em qualquer reunião de adolescentes. (“Seductive girls are the center of attention in parties and nightclubs at any time. Marie is a young woman who charms boys in any teenage gathering.”)

7. *De tempos em tempos, animais adestrados ficam famosos em filmes de sucesso. Argento é um cavalo que estrela a história Cavaleiro Solitário na pele de Silver. (“From time to time, trained animals become famous in successful movies. Argento is a horse that stars in the Lone Ranger story playing the role of Silver.”)*

Em alguns lugares, maridos ciumentos podem agredir esposas sem punição. Anderson é um cavalo que recebeu até hoje apenas advertências por bater na mulher. (“In some places, jealous husbands may beat their wives without any punishment. Anderson is a jackass who has only received warnings for beating his wife.”)*

Em alguns lugares, maridos ciumentos podem agredir esposas sem punição. Anderson é um marido que recebeu até hoje apenas advertências por bater na mulher. (“In some places, jealous husbands may beat their wives without any punishment. Anderson is a husband who has only received warnings for beating his wife.”)

8. *Muito raramente, heroínas de filmes são veículos a motor femininos. Rochelle é um avião que atua no sucesso de animação da Disney chamado Aviões. (“Movie heroes are seldom female motor vehicles. Rochelle is an airplane who acts in the Disney animation hit “Planes.””)*

De modo esperado, modelos de beleza são convidadas a posar para fotos. Rochelle é um avião que desfila com sucesso nas capas de revistas e nas passarelas. (“As expected, beauty models are invited to pose for photos. Rochelle is a beauty who appears successfully in magazine covers and runways.”)*

De modo esperado, modelos de beleza são convidadas a posar para fotos. Rochelle é uma jovem que desfila com sucesso nas capas de revistas e nas passarelas. (“As expected, beauty models are invited to pose for photos. Rochelle is a young woman who appears successfully in magazine covers and runways.”)*

9. *Para quase todos, produtos japoneses são tidos como muito bons e duráveis. Tagima é um violão que músicos destacados têm como referência de qualidade. (“For almost everyone, Japanese products are viewed as good and durable. Tagima is a guitar that successful musicians use as reference for quality.”)*

De modo habitual, garotas de biquíni são criticadas com rigor por mulheres. Tagima é um violão que revela destemor diante dos olhares de todas na praia. (“Girls in bikinis are usually criticized by women. Tagima is a curvaceous woman who does not fear being seen by others on the beach.”)*

De modo habitual, garotas de biquíni são criticadas com rigor por mulheres. Tagima é uma moça que revela destemor diante dos olhares de todas na praia. (“Girls in bikinis are usually criticized by women. Tagima is a young woman who does not fear being seen by others on the beach.”)*

A técnica de ERP no processamento de sentenças de crianças: uma revisão

The ERP technique in children's sentence processing: a review

Marília Uchôa Cavalcanti Lott de Moraes Costa

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil

marilia@letras.ufrj.br

Resumo: A medição da ativação cerebral por meio da técnica de potenciais relacionados a eventos (ERP) tem sido valiosa para lançar luz sobre diversas cognições humanas. A linguagem é uma das cognições que têm sido estudadas com essa técnica de grande resolução temporal entre o estímulo apresentado e a ativação observada decorrente desse estímulo. A área de aquisição da linguagem tem se beneficiado especialmente dessa técnica, dado que é possível investigar relações entre dados linguísticos e a ativação cerebral sem a necessidade de uma resposta explícita, como apertar um botão ou apontar para uma imagem. O objetivo deste artigo é apresentar o estado da arte sobre o processamento de frases em crianças utilizando a técnica de ERP.

Palavras-chave: ERP/EEG; processamento de frases; aquisição da linguagem; N400; P600.

Abstract: The measurement of brain activation through the technique of event related potentials (ERP) has been valuable in shedding light on various human cognitions. Language is one of those cognitions that has been studied with this technique, which allows for more accurate temporal resolution between the stimulus presented and the time in which we observe an activation resulting from this stimulus. The area of language acquisition has especially benefited from this technique since it

is possible to investigate relationships between linguistic data and brain activation without the need for an explicit response. The purpose of this article is to present the state of the art in children's sentence processing captured by the ERP technique.

Keywords: EEG/ERP; sentence processing; language acquisition; N400; P600.

Recebido em: 10 de dezembro de 2016.

Aprovado em: 4 de abril de 2017.

O objetivo deste artigo é revisar a literatura que utiliza a técnica de extração de potenciais elétricos de um traçado de eletroencefalograma (EEG/ERP) no campo de processamento de sentenças em crianças. Essa técnica tem sido utilizada por diferentes laboratórios, para medir diferentes modos de ativação cerebral relacionados a tarefas cognitivas e já demonstrou resultados importantes. Neste artigo, pretendo demonstrar resultados que atestam que crianças de 18 meses tem sensibilidade quanto à categoria de palavras e com 24 meses distinguem frases mal formadas daquelas bem formadas na língua que estão adquirindo. Estudos com crianças de até 4 anos são apresentados demonstrando a vasta competência linguística de crianças nessa faixa etária. Esse artigo também apresenta a técnica de ERP e os principais componentes encontrados em adultos, dessa forma é possível perceber as diferenças e semelhanças entre o processamento adulto e o infantil. Essa revisão procura contribuir com a área de aquisição da linguagem e neurociência da linguagem trazendo o estado da arte nessas duas áreas para o público brasileiro.

Para isso é relevante situar brevemente o leitor a respeito da técnica EEG-ERP. A eletroencefalografia (EEG) é uma técnica de imagem que mensura a atividade elétrica gerada por estruturas cerebrais (TEPLAN, 2002) na superfície do escalpo. Voltagens do escalpo refletem diretamente atividade neuronal. O EEG refere-se à atividade elétrica gravada continuamente da superfície do escalpo após ser capturada por canais, eletrodos de metal. O EEG é a mais antiga das técnicas de imagem. Hans Berger reportou a primeira gravação em 1929 (cf. LUCK, 2005, para informações históricas sobre a técnica).

O EEG tem uma resolução temporal muito boa, na ordem de alguns milissegundos; ele é um procedimento não invasivo que pode ser utilizado tanto em adultos como em crianças. É relativamente barato quando comparado a outros sistemas, como a ressonância magnética funcional (fMRI) e o PET (Tomografia por emissão de Pósitrons), métodos hemodinâmicos, ou o magneto-encefalograma (MEG), e ele não requer pessoal especializado da área médica. Outro aspecto importante quando comparamos o EEG com métodos hemodinâmicos é que, diferente desses, o ambiente de testagem do EEG é silencioso, o que é ideal para estudos de linguagem que utilizam estímulos auditivos.

O EEG tem a desvantagem de ter uma menor resolução espacial, dado que, a eletricidade captada é gerada dentro do crânio e perpassa diversas barreiras como diferentes tipos de tecido cerebral e barreiras celulares até alcançar o escalpo. O ricochetear que a atividade elétrica sofre até chegar à superfície faz com que a ativação captada no escalpo e o local onde ela foi gerada não sejam análogos, ou seja, a fonte da atividade elétrica não está normalmente localizada logo abaixo do eletrodo onde está sendo feita a mensuração.

No sinal contínuo gravado pelo EEG, estão as respostas neuronais associadas a eventos sensoriais, linguísticos, e motores – assim como atividade neuronal não relacionada e ruídos (TEPLAN, 2002). Para analisar as respostas elétricas correspondentes a um evento específico, a técnica de potenciais relacionados a eventos é utilizada (JOHNSON, 2001; LUCK, 2005). Essa técnica consiste em tirar uma média da atividade cerebral de muitos *trials* (item de um experimento) que correspondem à mesma condição experimental. A média permite que os pesquisadores eliminem o ruído e as atividades neuronais não relacionadas ao experimento.¹ Ao calcular a média do sinal, assume-se que o sinal de interesse ocorra de maneira consistente através dos *trials* enquanto que o ruído é aleatório (LUCK, 2005).

No entanto, para garantir a exclusão do ruído indesejado e não do sinal relevante é necessário ter um grande número de itens experimentais para realizar as médias, termo conhecido como razão sinal-ruído ou relação sinal-ruído. Quanto maior o sinal, menor é a influência do ruído. Isso quer dizer que para obter uma onda ERP confiável em um experimento de linguagem padrão, a média deve ser feita com pelo

¹ Para facilitar a leitura chamarei toda ativação não relacionada ao experimento de ruído.

menos 20-30 *tokens* diferentes dentro da condição experimental. Assim, quando falamos do ERP suscitado por uma palavra específica em uma condição particular, queremos dizer a atividade eletrofisiológica média em diferentes *tokens* do mesmo tipo (BROWN; HAGOORT; KUTAS, 2000).

O método de ERP pode identificar diretamente o potencial relacionado ao processamento de um estímulo. Ele é muito utilizado em estudos de aquisição da linguagem (OBERECKER *et al.*, 2005; OBERECKER; FRIEDERICI, 2006; SILVA-PEREYRA *et al.*, 2005a, 2005b; SILVA-PEREYRA *et al.*, 2007; BERNAL *et al.*, 2010; BRUSINI *et al.*, 2016a, 2016b; COSTA, 2015) e processamento linguístico em geral (KUTAS *et al.*, 1980; FRANÇA *et al.*, 2004; GOMES; FRANÇA, 2013; GOMES, 2014; GOMES; FRANÇA, 2015). Outra vantagem de ERPs é que não há a necessidade de os participantes fazerem uma tarefa – o que torna essa tecnologia mais atrativa em populações mais jovens que podem participar de estudos de linguagem sem receber instruções (CONBOY *et al.*, 2008). Na maioria dos estudos com ERP os voluntários recebem instruções para realizar uma tarefa apenas para mantê-los atentos, mas a atividade cerebral pode ser capturada sem nenhuma tarefa específica.

Após décadas de estudo, alguns componentes elétricos foram frequentemente observados – N170, ELAN, N400, P300, P600 – quando participantes processam estímulos específicos (como fala, faces, etc) e hoje eles são considerados marcadores de certos tipos de processamento realizado pelo cérebro. Há restrições para qualificar uma certa atividade cerebral encontrada por um paradigma experimental como um componente elétrico já muito conhecido. Existem critérios bem definidos para que uma dada ativação cerebral seja caracterizada como um N400 ou P600, por exemplo. A população de neurônios envolvida na atividade identificada precisa ser grande o suficiente e a atividade deles precisa ser sincronizada e coerente durante uma janela de tempo com duração suficiente; em nosso exemplo, a duração mínima seria 200ms. A definição de um potencial consiste de três dimensões da atividade gravada:

1. Latência (ms): o tempo em que um potencial começa, se desenvolve e desaparece;
2. Distribuição Espacial: a disposição dos eletrodos que exibem o efeito;
3. Voltagem: a polaridade do efeito (positiva ou negativa).

Esses componentes são comumente nomeados levando em consideração a polaridade deles (N ou P) e o número de milissegundos em que se observa tipicamente o efeito máximo (e.g.: N400, P600).

A literatura sobre aspectos linguísticos no cérebro já reúne um número considerável de trabalhos e há muitas hipóteses sobre como se dá o processamento linguístico. Poeppel e Embick (2005) e Marantz (2005) expressam as dificuldades e a importância da interseção entre a Linguística e a Neurociência (POEPPEL; EMBICK, 2005; MARANTZ, 2005). Desafios importantes que precisam de atenção são chamados por Poeppel e Embick de *incompatibilidade de granularidade e incomensurabilidade ontológica*. O primeiro conceito está relacionado com o fato de haver uma incongruência entre os elementos representacionais das duas disciplinas. Em geral a linguística trabalha com elementos bem detalhados, traços distintivos, morfema, sintagmas etc., enquanto que em abordagens neurocientíficas os termos são vistos de maneira mais ampla, dendritos, neurônios, coluna cortical. O segundo conceito está relacionado com o fato desses elementos das teorias linguísticas não poderem ser diretamente comparados com as unidades observadas pela Neurociência. Os autores apontam que esses problemas não são específicos da área da linguagem, mas são “problemas de interface” comuns aos estudos da cognição” (POEPPEL; EMBICK, 2005).

Apesar de a possibilidade de descobrir os mecanismos subjacentes ao processamento linguísticos ser empolgante, os elementos que eliciam ou afetam algum componente de ERP não se equiparam a uma função neurocognitiva. Por enquanto, a forma mais direta de observar um resultado de ERP é apresentar eventos altamente controlados que variam em apenas uma variável e compará-los entre muitos sujeitos. Se uma diferença for encontrada, então deve ser pelo fato de os estímulos terem alguma propriedade que o cérebro dos participantes “pôde diferenciar”.

Cada estudo postula uma explicação que se encaixa aos dados e possivelmente a uma teoria, mas é apenas após um número de manifestações que seremos capazes de entender melhor o que os elementos descritos na literatura de ERP como componentes distintos quer dizer sobre os mecanismos subjacentes. A cada novo estudo entendemos mais sobre a linguagem, o cérebro e a técnica. Muitos anos após o primeiro estudo de ERP em linguagem, o debate ainda está aberto sobre o que significam para a linguagem os vários componentes de ERP encontrados e como integrar esses estudos com o corpo massivo de

trabalhos já existentes em psicolinguística para alcançarmos os melhores modelos sobre processamento de linguagem. Como Pylkkänen e colegas afirmam:

In this endeavor, a cognitive model is crucial as without it, the cognitive neuroscientist does not know what to look for in the brain, what the nature of the relevant representations might be, or how the different components of a process might interact with each other.” (PYLKKÄNEN; BRENNAN; BEMIS, 2011).

Componentes comumente encontrados em experimentos de processamento de frases em adultos

Muito já foi aprendido sobre o fato de anomalias linguísticas nos níveis semântico, morfológico e sintático eliciarem uma série de respostas de ERP com características diferentes. Estudos com anomalias sintáticas e morfológicas, por exemplo, têm demonstrado respostas qualitativamente diferentes de ERP aos diferentes tipos de violações. Abaixo faremos uma revisão destes tipos de componentes.

Como já foi mencionado, componentes de ERP são normalmente denominados de acordo com sua polaridade e latência. No entanto alguns componentes podem ser nomeados a partir de sua topografia. LAN (*Left Anterior Negativity*, negatividade anterior à esquerda) e ELAN (*Early Left Anterior Negativity* – negatividade anterior à esquerda precoce) são dois exemplos na área da linguagem. As duas ondas aparecem cedo e são eliciadas, de acordo com a literatura, por violações sintáticas. LAN e ELAN são, como o nome sugere, ondas negativas capturadas à esquerda nas áreas anteriores do cérebro. Essas duas ondas, analisadas como idênticas, diferem apenas em sua latência. O ELAN é muito precoce e aparece entre 150 e 200 milissegundos após o começo do elemento crítico; já o LAN se desenvolve entre 300-500 milissegundos. Estudos reportaram que o LAN e o ELAN opunham frases neutras, gramaticais, de frases contendo violações sintáticas: a morfossintaxe não estava adequada (concordância ou marcador de caso repetido) ou em sentenças contendo uma combinação de palavras inadequadas (e.g. uma palavra ou grupo de palavras aparecia em um contexto em que palavras de uma categoria sintática diferente eram esperadas – HAGOORT; LEVELT, 2009; HWANG; STEINHAUER, 2011).

O N400

No final dos anos 70, Martha Kutas e Steven Hillyard estavam investigando a influência do contexto sentencial para o reconhecimento de palavras. Eles inicialmente pensaram que a manipulação do grau de previsibilidade do final de frases os levaria a observar o componente P3², que é normalmente relacionado à probabilidade de um certo estímulo aparecer. Os autores manipularam o final de frases, que variavam de congruentes até ininterruptas, mas, ao invés de uma onda posterior positiva, eles encontraram uma onda negativa. Na verdade, a única manipulação que deu o gatilho para um P3 foi o uso de finais congruentes, porém com fontes distintas, letras maiúsculas, na última palavra. Essa nova forma de onda foi chamada de N400 devido à sua polaridade e ao fato de seu pico ser observado em torno dos 400ms (KUTAS; HILLYARD, 1980). Em 1980, o primeiro artigo sobre o N400 foi publicado. Um par de exemplos desse trabalho foi “*I like my coffee with cream and sugar.*” versus “*I like my coffee with cream and socks.*”³. O segundo tipo de frase levaria a uma onda negativa 400 ms após o elemento crítico que nesse caso é a palavra final (*sugar* ou *socks*). Esse trabalho seminal foi feito com estímulos visuais, escritos. Naquele momento, os autores interpretaram a onda como uma resultante de uma violação semântica. Um grande número de trabalhos replicaram esse efeito estendendo os achados a outras modalidades (HOLCOMB; COFFEY; NEVILLE, 1992) e outros contextos linguísticos, incluindo pares de palavras e contextos discursivos mais amplos, assim como sentenças (OSTERHOUT; HOLCOMB, 1992; BERKUM; HAGOORT; BROWN, 1999; FRANÇA, 2002; FRANÇA *et al.*, 2004; HÄNEL-FAULHABER *et al.*, 2014).

Hoje, o N400 é descrito como uma deflexão negativa ampla do ERP que começa 200–300 ms após uma palavra ou sinal, no caso das línguas de sinais, ser apresentada por via auditiva ou visual e que tem seu pico depois de 400ms, aproximadamente, em eletrodos centro-parietais (cf. KUTAS; FEDERMEIER, 2011; LAU; PHILLIPS; POEPEL, 2008). Pesquisadores observaram o N400 em um número variado de tarefas que variam de *priming* de palavras e processamento no nível

² Uma onda positiva que tem seu pico normalmente 300 milissegundos após o elemento crítico referida como P3.

³ A primeira frase lê-se “Eu gosto do meu café com creme e açúcar” e a segunda “Eu gosto do meu café com creme e meias”

da mensagem até o papel da atenção e consciência na compreensão da linguagem, diferenças de ativação entre hemisférios ativadas pela memória semântica, e a construção de vocabulário no aprendizado de uma língua (OSTERHOUT; MCLAUGHLIN, 2006; KOTZ; HOLCOMB; OSTERHOUT, 2008), só para mencionar alguns trabalhos. O N400 também já foi usado para comparar diferenças entre populações – e.g., jovens vs. idosos, sadios vs. pacientes esquizofrênicos (KUPERBERG *et al.*, 2006).

In general then, the N400 is thought to reflect the degree of ease or difficulty in retrieving stored knowledge associated with a potentially meaningful item from semantic memory, contingent upon both the characteristics of the stored item itself, as well as the contextual cues available. (KUTAS; DELONG, 2008, p. 157)

Apesar de uma resposta N400 ser majoritariamente associada a uma anomalia semântica, ela também já foi eliciada pela maioria de estímulos significativos – palavras faladas, escritas ou sinalizadas, não-palavras pronunciáveis, e acrônimos conhecidos, mas também para sons não-humanos, faces, gestos, e cenas. Desde a primeira vez que a onda foi reportada, pesquisadores têm tentado explicar os mecanismos subjacentes relacionados ao N400. O componente já foi estudado a partir de diferentes ângulos e campos de investigação e mesmo assim ainda hoje há muitos aspectos que não são completamente compreendidos. De acordo com Kutas e Federmeier (2011), um ponto crucial da literatura de N400 é o panorama convincente de como a percepção, a atenção, a memória, e a linguagem em conjunto participam nos eventos neuronais responsáveis pelo N400 (KUTAS; FEDERMEIER, 2011). Apesar de não ser eliciado exclusivamente em estudos de processamento de linguagem, o N400 é uma ferramenta para investigar questões relacionadas ao curso temporal do processamento de linguagem. Um ponto permanece em discussão: exatamente qual(is) processo(s) o componente reflete?

O P600

O P600 foi descrito pela primeira vez em 1992 por Osterhout e Holcomb. É uma onda com voltagem positiva que apresenta sua amplitude máxima em torno de 600ms após o estímulo crítico. Ele é eliciado tanto

na modalidade visual quanto na auditiva. Ativa tipicamente eletrodos centro-parietais, mas alguns trabalhos já reportaram esse componente em eletrodos frontais (FRIEDERICI, 2002). O P600, também chamado de *Syntactic Positive Shift* – mudança sintática positiva (SPS), devido a sua natureza positiva e ao fato de ser eliciado por estímulos sintáticos (OSTERHOUT; HOLCOMB, 1992; FRIEDERICI; PFEIFER; HAHNE, 1993; HAGOORT; BROWN; GROOTHUSEN, 1993; OSTERHOUT; HOLCOMB; SWINNEY, 1994; GOMES; FRANÇA, 2013; GOMES, 2014; GOMES; FRANÇA, 2015).

Originalmente, o P600 foi interpretado como um reflexo de processos sintáticos. Ele foi observado em resposta a diferentes violações sintáticas, incluindo violações na estrutura sintagmática (OSTERHOUT; HOLCOMB, 1992; HAGOORT; BROWN; GROOTHUSEN, 1993), violações de subcategorização (OSTERHOUT; HOLCOMB, 1992; OSTERHOUT; HOLCOMB; SWINNEY, 1994), violações quanto ao número, tempo, gênero, e concordância de caso (COULSON; KING; KUTAS, 1998; NEVINS *et al.*, 2007) e violações às restrições de dependência de longa distância (PHILLIPS; KAZANINA; ABADA, 2005). Devido à variedade de circunstâncias em que o P600 foi observado, ele foi interpretado como o reflexo de reparo sintático – e.g., (OSTERHOUT; HOLCOMB, 1992; FRIEDERICI, 1995), reanálise estrutural, ou integração sintática (GIBSON, 1998; KAAAN *et al.*, 2000; HERTEN; KOLK; CHWILLA, 2005). No entanto, nos últimos 10 anos a interpretação de que o P600 é especificamente relacionado a processamento sintático tem sido questionada (KUPERBERG; SITNIKOVA, 2003; KOLK *et al.*, 2003; KIM; OSTERHOUT, 2005). Hoje já se discute o P600 como um componente que pode refletir tanto o componente sintático como semântico (BORNKESSEL-SCHLESEWSKY; SCHLESEWSKY, 2008; HOEKS; STOWE; DOEDENS, 2004; KUPERBERG, 2007).

Gouvea *et al.* (2010) estudaram todos os “estímulos P600” comumente descritos para um melhor entendimento da natureza dessa onda. Eles utilizaram um desenho experimental intra-sujeito com o máximo de estímulos similares para três tipos de configuração estruturais que havia sido discutido na literatura como “eliciadores de P600” – sentenças agramaticais, *garden paths*, e complexidade de dependências de longa-distância. O resultado foi que a topografia para *garden paths* sintáticos, violações de gramaticalidade, e preenchimento de dependências de longa-distância eram bastante semelhantes, mas

as dependências com relação a palavras qu- tinham uma distribuição diferente, inicialmente mais anterior. A condição das dependências de qu- eliciaram um P600 de amplitude menor, e as condições com frases *garden path* e agramaticais apresentaram uma negatividade adicional e “o P600 posterior tinha uma latência retardada nas condições agramaticais (500-700 ms) do que nas condições com frases *garden path* (300-500 ms)” (GOUVEA *et al.*, 2010).

Gouvea *et al.* (2010) sugerem que o P600 possa refletir um conjunto comum de processos que começam assim que informações suficientes foram acumuladas para iniciá-los. Sob essa visão, mesmo quando o P600 ocorre em latências distintas, ele ainda reflete os mesmos processos subjacentes, a diferença de latência refletiria mais ou menos o tempo necessário para completar os processos que acionam o P600.

Uma vez que a positividade que atinge picos de 600 ms é provocada por diferentes tipos de estímulos, ela levanta naturalmente a questão de se ela reflete um único mecanismo subjacente ou vários processos diferentes. Por enquanto, alguns pesquisadores discutiram possíveis maneiras de explicar o P600 como refletindo diferenças nos cálculos realizados por um único processo subjacente, mas ainda pode haver a possibilidade de que pequenas diferenças de topografia, tensão e tempo de pico sejam índices de diferentes mecanismos subjacentes.

Estudos de linguagem ERP sobre processamento de sentenças em crianças pequenas

Não há muitos estudos de ERP testando processamento de sentenças em crianças, por várias razões: (i) artefatos de movimento; (ii) dificuldade de colocação dos eletrodos/ touca de EEG; (iii) participação dos responsáveis; (iv) duração dos experimentos e dificuldade de encontrar estímulos atraentes. Como foi explicado anteriormente, uma das principais dificuldades para uma gravação de qualidade é o movimento. Solicita-se aos participantes adultos que não se movam e se abstenham de piscar em partes críticas do teste. Em contraste, você não pode pedir a uma criança pequena para ficar quieta, especialmente quando ela tem entre 14 a 36 meses – a idade exata quando você precisa estudar o início do processamento da sentença. É muito cansativo para eles permanecer no lugar por um longo tempo. A última barreira é a duração dos experimentos: embora o EEG seja a técnica ótima para estudos de

crianças porque é silencioso e as crianças não têm que dar respostas abertas, o ERP exige a média de múltiplos *trials* em cada condição, o que torna as experiências geralmente muito longas para crianças (20-30 minutos). Se uma criança não participar do experimento por um período de tempo suficiente, o número de itens será muito baixo para obter médias estáveis. Para incentivar as crianças a permanecerem no experimento tempo suficiente, os estímulos precisam ser muito atraentes para elas. Todas essas barreiras podem explicar por que há tão poucos estudos com crianças pequenas na literatura ERP.

Uma última dificuldade com as crianças é que um cérebro em desenvolvimento não tem as mesmas propriedades de um cérebro totalmente desenvolvido⁴ (PUJOL *et al.*, 2006). Não é raro que os componentes de ERP de crianças sejam diferentes temporalmente, topograficamente ou quanto à polaridade. É difícil, portanto, usar o conhecimento que temos da literatura adulta para extrapolar o que deve ser encontrado em populações jovens (COCH, MITRA, GEORGE, 2012). As diferenças podem ocorrer em latência, duração e amplitude.

Embora haja muitos obstáculos na pesquisa de desenvolvimento de ERP, ainda existem alguns grupos ao redor do mundo que realizam essa empreitada. Encontramos estudos de ERP em diferentes domínios cognitivos e idades (KOVELMAN, 2012). No estudo da língua, a maioria dos estudos é feita em inglês, francês ou alemão. Existem muitos estudos em prosódia e fonologia (MILLS *et al.*, 2004). Há evidências de que as crianças podem discriminar entre categorias de som, padrões de tonicidade (*stress*) (FRIEDERICI; FRIEDRICH; CHRISTOPHE, 2007) e unidades prosódicas, por exemplo. Os estudos em fonologia geralmente são realizados com bebês bem pequenos, que não se importam com as toucas; em muitos casos, os estudos podem ser feitos até mesmo quando o bebê está desatento à tarefa (por exemplo, distraído com bolhas de sabão enquanto ouve sílabas ou palavras) e às vezes até durante o sono (DEHAENE-LAMBERTZ; PENA, 2001).

Há também trabalhos sobre a segmentação da fala. O trabalho realizado com crianças de até 7 meses de idade apresentou evidência

⁴ Cf. (SHAFER; GARRIDO-NAG, 2007) para uma revisão sobre as evidências neurobiológicas relativas ao desenvolvimento da linguagem e (POEPEL; OMAKI, 2008) para uma discussão sobre as ressalvas e perspectivas do uso de abordagens ERP no desenvolvimento da linguagem.

significativa de segmentação da fala. Os participantes mostraram uma polaridade positiva (KOOIJMAN *et al.*, 2013). Os estudos foram conduzidos com técnica ERP para investigar o acesso lexical também. Os pesquisadores desenvolveram um paradigma apropriado para adultos e crianças, no qual o participante vê uma imagem de um objeto e ao mesmo tempo ouve um estímulo auditivo que é uma palavra correspondente ao nome do objeto ou não, ou é uma pseudopalavra que é fonotaticamente bem ou mal formada – em português poderíamos ter exemplos como baluga e btínica, respectivamente. Usando esse paradigma, Friedrich e Friederici (2004) observaram uma mudança de desenvolvimento entre as idades de 12 e 19 meses. Os efeitos de ERP em crianças de 19 meses foram relatados como bastante semelhantes aos dos adultos, isto é, um efeito N400 para as palavras incongruentes e pseudopalavras fonotaticamente possíveis, mas não para pseudopalavras fonotaticamente ilegais (FRIEDRICH, FRIEDERICI, 2004). No entanto, eles não encontraram uma negatividade aos 400 ms para 12 meses de idade. Diferentemente dos resultados encontrados com 12 meses de idade, em outro estudo realizado em 2010, os autores observaram um efeito N400 de *priming* com a mesma população. A ativação foi relacionada ao nível de produção de palavras de bebês. Os participantes foram divididos em dois grupos: um com alta produção de palavras precoces, e o outro com baixa produção de palavras precoces. Apenas o primeiro grupo teve o que Friedrich e Friederici interpretou como um N400 (FRIEDRICH; FRIEDERICI, 2010).

De acordo com Silva-Pereyra e colegas (2005b), a maior parte das crianças apresenta dois momentos de desenvolvimento sem igual após o primeiro ano de vida. O primeiro é uma aceleração na velocidade do crescimento do vocabulário a partir de 16 e 20 meses (SILVA-PEREYRA *et al.*, 2005b). Aos 20 meses de idade, diferenças de ERP entre palavras conhecidas e não conhecidas se tornam lateralizadas e mais diferenciadas nas regiões temporal e parietal do hemisfério esquerdo, possivelmente refletindo o aumento de especialização dessas áreas para o processamento da linguagem (cf. os trabalhos do grupo de Debra Mills para uma revisão desses resultados). O segundo é no desenvolvimento da morfossintaxe das crianças, entre 24 e 30 meses de idade. Esse segundo momento é mais relevante para esse trabalho, dada a relação com o processamento de sentenças.

Oberecker *et al.* (2005) e Oberecker e Friederici (2006) trabalharam no processamento de sentenças em crianças de 2 anos, de 28 meses e de 24 meses de idade, respectivamente. O estudo foi realizado em alemão. Eles testaram se nessa idade as crianças eram sensíveis às estruturas sintáticas de sua língua, como exemplificadas em (1), (2) e (3):

- (1) Der Löwe brüllt
O leão ruge (ou está rugindo)
- (2) Der Löwe im Zoo brüllt
O Leão no Zoo rugindo
- (3) Der Löwe im brüllt
O leão no rugindo

As sentenças foram apresentadas a crianças em três condições – duas condições gramaticais e uma agramatical. As frases gramaticais tinham ou estrutura de DP (VP) como em (1) ou estrutura de DP [PP VP] como em (2). A condição agramatical tinha sentenças com descasamento de maiúsculas e minúsculas; o caso era marcado por uma preposição que tomava um nome ou adjetivo, mas as frases tinham um verbo em vez disso como em (3). Os resultados apresentados em Oberecker, Friedrich, Friederici (2005) demonstraram que os participantes de 2 anos e 8 meses exibiram o que os autores chamaram de ELAN e P600, enquanto que aqueles de 2 anos exibiram apenas um P600 (OBERECKER; FRIEDERICI, 2006). O padrão encontrado no grupo mais velho estava alinhado com o grupo-controle adulto que também exibiu um ELAN e P600. Os autores interpretaram esses achados sugerindo que os princípios de sintaxe são estabelecidos precocemente durante o desenvolvimento, uma vez que um P600 estava presente no grupo mais jovem, pelo menos no que se refere a frases simples.

According to this interpretation the presence of a P600 effect in 2-year-olds would mean that late integration processes are already at work at this age even though highly automatic syntactic processes reflected in the ELAN are not yet effective. The biphasic ELAN-P600 pattern observed in 2.8-year-olds indicates that local phrase

structure building processes are already established by that age in addition to late integration processes at least for the processing of simple active sentences. A similar developmental shift from a P600-only to a biphasic ELAN-P600 pattern was reported for the processing of phrase structure violations in passive sentences between the age of 6 and 7 years (Hahne, Eckstein, and Friederici, 2004). (FRIEDERICI; THIERRY, 2008, p. 227)

Em 2010, Bernal e colegas estudaram se as crianças construíam expectativas online em relação à categoria sintática da próxima palavra em uma frase. Os pesquisadores apresentaram frases que tinham substantivos no lugar de verbos e vice-versa. Os resultados mostraram respostas cerebrais diferentes para sentenças gramaticais e não-gramaticais, sugerindo que aos 24 meses de idade, as crianças constroem expectativas on-line sobre a categoria das palavras. Outro achado interessante foi que a topografia da resposta cerebral foi diferente para substantivos e verbos, sugerindo que o processamento de substantivos e verbos em crianças pequenas já poderia ser conduzido por mecanismos subjacentes diferentes (como é encontrado em adultos).

Em 2012, o mesmo grupo relatou resultados semelhantes com novas palavras (BRUSINI, 2012 – tese de doutorado). Eles ensinaram às crianças substantivos e verbos novos⁵ criados para o teste com a fonotática do francês, *touse* como em (4), uma semana antes do teste – nunca com a estrutura alvo –, e, no dia do teste, como em Bernal *et al.* (2010), eles os apresentaram em frases gramaticais como em (5) ou agramaticais nas quais os substantivos ensinados na semana anterior eram inseridos agora em estrutura verbal, como em (6), e vice-versa.

(4) Oh regarde, un touse!

Ih Olha, um touse!

(5) L'indien pousse le [touse]_N

O índio empurra o *touse*.

⁵ As palavras inventadas eram ensinadas a partir de brincadeiras com bonecos. *Touse*, por exemplo, foi ensinado com o significado de tricerátops.

(6) *Alors, elle le [touse]_v de joie.

Então, ela o touse de alegria.

Os pesquisadores também encontraram respostas diferentes para frases gramaticais e agramaticais, embora com uma topografia e uma janela temporal diferentes do que aquelas do estudo de 2010. Essa tese se converteu em um número de artigos publicados em 2016 – BRUSINI *et al.*, 2016a, 2016b).

Ainda em relação ao processamento sintático, e mais especificamente aos verbos e à morfossintaxe, em 2005, dois artigos do mesmo grupo de pesquisadores surgiram. O estudo com crianças mais velhas foi publicado em primeiro lugar e, em seguida, um grupo mais jovem também foi testado e os resultados desses testes foram publicados em um artigo separado. Isso significa que, uma vez que os autores observaram os potenciais obtidos pela manipulação de propriedades morfossintáticas e semânticas, eles estavam confiantes de que tal ativação poderia ser encontrada em crianças mais novas. É comum na literatura do ERP de desenvolvimento procurar primeiro os fenômenos em um grupo etário e, em seguida, tentar observar quão cedo pode-se observar o fenômeno. Silva-Pereyra *et al.* (2005b) abordaram o curso temporal dos ERPs utilizando frases sintáticas e semanticamente anômalas para comparar com as controle, com crianças de 36 meses de idade (SILVA-PEREYRA; RIVERA-GAXIOLA; KUHL, 2005b). O objetivo era observar que tipo de ativação cerebral seria observada quando sentenças sintática e semanticamente anômalas foram apresentadas aos participantes em idade pré-escolar. Silva-Pereyra *et al.* (2005a) apresentaram o mesmo conjunto de estímulos em uma população mais jovem, de 30 meses de idade (SILVA-PEREYRA *et al.*, 2005a).

Silva-Pereyra e colaboradores estudaram crianças de 36 e 48 meses, manipulando elementos sintáticos e semânticos (SILVA-PEREYRA, RIVERA-GAXIOLA, KUHL, 2005). Eles adicionaram o sufixo *-ing* às estruturas em que ele não é licenciado como em (7) e alteraram o argumento interno dos verbos tornando-os agramaticais, como em (8).

(7) My uncle will watching a movie.

IPOS tio FUT assistir-ASP DET filme

Meu tio irá assistindo um filme.

(8) My uncle will blow the movie.

1POS tio FUT soprar DET filme

Meu tio vai soprar um filme.

Com essa manipulação, os autores demonstraram que crianças de 36 e 48 meses apresentaram efeitos ERP diferentes para processamento sintático e semântico (mesmo sem atenção explícita às frases) e que suas respostas mudaram entre 3 e 4 anos de idade. As ativações tiveram localização mais anterior aos 3 do que aos 4 anos de idade. O mesmo padrão foi replicado em crianças de 2,5 anos de idade.

Eles descobriram uma mudança positiva significativa na forma de onda durante o processamento de frases sintaticamente anômalas quando comparadas com frases não anômalas. Este efeito foi composto apenas de um efeito positivo duradouro que foi maior sobre as regiões anteriores do escalpo de 500 a 1500 ms. A interpretação foi que o efeito foi consistente com os resultados de P600 de adultos durante o processamento sintático. Os P600 adultos sob o mesmo tipo de manipulação são apresentados em eletrodos de linha média parietal (HAGOORT; BROWN; OSTERHOUT, 1999). Entretanto, o efeito sintático em pré-escolares tinha uma distribuição anterior. Uma possível conclusão após a análise desses resultados é que, já em tenra idade, as crianças têm a capacidade de analisar a sintaxe. Eles implementam diferentes processos ao analisar sentenças gramaticais e não-gramaticais. Em geral, o potencial evocado foi uma onda de positividade tardia. Este resultado foi interpretado como semelhante ao P600 adulto, mas os pré-escolares tiveram uma resposta mais distribuída nos eletrodos anteriores. O mesmo padrão estava presente em crianças de 2,5 anos (SILVA-PEREYRA *et al.*, 2005a). A conclusão geral foi de que as crianças pequenas possuem assinaturas neurais do processamento das sentenças no início de seu desenvolvimento linguístico.

Em outro estudo dos mesmos autores (SILVA-PEREYRA *et al.*, 2005b) são comparadas as ativações de crianças de 36 e 48 meses em frases com violações morfossintáticas e semânticas. O estudo demonstrou que crianças de 36 meses e 48 meses demonstram efeitos de ERP diferentes para processamento sintático e semântico sem atenção explícita às sentenças e que suas respostas mudam entre 3 e 4 anos de idade para as violações sintáticas. Uma onda positiva foi observada

nas frases sintaticamente anômalas nos participantes de 36 meses com pico de ativação em 800 ms e duas ondas positivas na população de 48 meses: uma com pico aos 400 ms e a outra em 800 ms. Essas ativações foram reportadas como P400 e P800. Os autores também reportaram uma diferença quanto à topografia, que demonstrou ativações mais fortes em regiões frontocentrais. Em relação às frases com anomalias semânticas, três ativações negativas foram reportadas como N400, N600 e N800, nos dois grupos.

Outra área de investigação na literatura da linguística e de processamento de frases é a importância das pistas sintáticas e semânticas. Essa importante questão também foi investigada em crianças. Para poder observar se as sugestões semânticas são mais salientes do que a estrutura das frases, Silva-Pereyra e colegas apresentaram a crianças de 36 meses com frases em inglês e em *jabberwocky*,⁶ sem conteúdo semântico – ou seja, usando apenas a estrutura de frases com palavras sem qualquer significado (SILVA-PEREYRA *et al.*, 2007). Eles testaram 16 crianças. Metade das sentenças eram agramaticais. As frases usadas tinham verbos transitivos como em (9). No exemplo, o argumento interno (*the cat*) é um DP. Os pesquisadores substituíram o tipo correto de argumento interno, um DP, por um sintagma preposicional como em (10). Esta é uma violação da estrutura do verbo “*touch*”. Nas frases *jabberwocky* a posição das palavras funcionais é respeitada em frases gramaticais como em (11) e (12), mas as palavras lexicais são substituídas por palavras possíveis mas inexistentes na língua.

(9) My dolly touched the cat with her hand.

‘Minha bonequinha tocou o gato com a mão dela.’

(10) My dolly touched with the cat her hand.

‘Minha bonequinha tocou com o gato a mão dela.’

(11) My cholly daunched the glat with her shond.

‘Minha tonequinha docou o clato com a prão dela.’

⁶ Jabberwocky é o termo para denominar sentenças com estruturas aparentemente perfeitas na língua mas que não fazem sentido algum. O termo é original do livro *Through the Looking-Glass, and What Alice Found There*, de Lewis Carrol (1871).

(12) My cholly daunched with the glat her shond

‘Minha tonequinha docou com o clato a prão dela.’

Se as crianças utilizassem o conhecimento semântico para determinar se uma frase era incongruente ou não, mais do que o conhecimento sintático, então seria de se esperar que os efeitos de ativação fossem mais fortes para frases que eram semanticamente incongruentes do que sintaticamente incongruentes. Os autores encontram ativações negativas e positivas maiores para as frases anômalas do que para as não-anômalas. No entanto, os resultados encontrados pelos autores são difíceis de interpretar dado que nenhum dos efeitos durou mais de 200ms. Na verdade, as ativações encontradas tinham apenas 150ms, o que é bastante rápido para esse tipo de fenômeno. Sendo assim, o que o leitor leu do artigo é a variedade de latência e localização que os autores encontraram que por vezes estavam alinhadas com outros estudos, mas também foram observadas latências e localização ainda não reportada.

O último estudo que será relatado diz respeito a uma investigação quanto à estrutura argumental. COSTA (2015) investigou se crianças de dois anos sabem inconscientemente que certos verbos entram em estruturas argumentais específicas como em (13), demonstrando efeito de agramaticalidade caso um verbo conhecido entrasse em outra estrutura como em (14), presente em outras línguas do mundo, mas não observada nos dados da língua francesa. Pequenos vídeos em que uma narradora contava uma história enquanto manipulava bonecos era apresentado às crianças. Ao longo das histórias, frases gramaticais – do tipo (13) e (15) e agramaticais como em (14) eram ouvidas e as pesquisadoras analisaram a ativação cerebral relacionada a essas frases. Os resultados demonstraram uma ativação negativa entre 200-450ms após a criança ter escutado o verbo que estava sendo manipulado. Esses resultados indicaram que as crianças são sensíveis a manipulações na estrutura argumental.

(13) Thomas lui donne une fleur pour lui dire merci.

‘Thomas lhe deu uma flor para lhe dizer obrigado.’

(14) * Elle lui donne qu’ils seront toujours amis.

‘Ela lhe deu que eles seriam sempre amigos.’

(15) Elle lui crie qu'il faut manger la bonne viande.

‘Ela lhe gritou que ele devia comer a boa carne.’

Todos esses estudos indicam que crianças bem pequenas que produzem palavras ou pequenas sentenças há poucos meses no momento dos testes já têm intuição linguística sobre o que é possível ou não em sua língua nativa quanto ao processamento de frases. Violações sintáticas e semânticas eliciam ondas diferentes daquelas encontradas em frases bem formadas. Os resultados apresentados aqui sugerem que crianças de um ano e meio, dois e três computam estruturas sintáticas complexas online. Isso nos leva a pensar sobre quais mecanismos já estão ativos nessa idade e a que outros tipos de subcategorização de verbos as crianças são sensíveis e quando essa(s) capacidade(s) estaria(m) disponível(is).

Conclusão

A ativação observada quando crianças processam sentenças não ocorre necessariamente na mesma topografia e janela temporal encontradas nos componentes clássicos de ERP. Alguns autores têm reportado ativações que ocorrem em janelas de 500-1500ms e em posições mais anteriores como P600. Isso estaria alinhado com o que Gouvea (2010) observa em sua revisão sobre o P600. No entanto, não sabemos dizer ainda se todos os elementos reportados como P600 em adultos vão ser encontrados na população infantil e se eles representarão os mesmos processos subjacentes. É importante dar continuidade às pesquisas para podermos, com um corpo maior de dados de outras línguas, conhecer mais sobre o processamento sintático infantil.

A área de estudo sobre desenvolvimento linguístico em ERP ainda é recente e as melhores formas de testar e interpretar os dados estão sendo continuamente debatidas, a fim de encontrar o melhor sistema. O campo está em ampla expansão e cada ano mais pesquisas e novos grupos são formados em todo o mundo. A comparação com adultos é necessária, mas às vezes, quando a ativação encontrada não está alinhada com aquela em adultos não há uma conclusão óbvia. A diferença pode existir devido ao nível de maturidade linguística ou fatores relacionados à anatomia infantil. Sendo assim, dizer apenas que os resultados são diferentes não necessariamente nos leva a conclusões claras sobre quais processos tal ativação indica.

Mesmo assim, o que sabemos hoje sobre o processamento infantil com ERP já é bastante encorajador. Crianças de 18 meses são sensíveis a manipulações quanto a categoria em palavras do dia-a-dia e aos 24 meses em palavras aprendidas no laboratório. Estudos mostram que, aos 24 meses, as crianças são sensíveis a manipulações quanto à marcação de caso e que aos 28 meses a ativação encontrada já apresenta padrão adulto. Aos 30 meses, as crianças são sensíveis a manipulações sintáticas e semânticas. Essa mesma manipulação elicia uma ativação com localização mais anterior em crianças de 3 do que em crianças de 4 anos. Pesquisadores também encontraram diferenças de ativação em crianças de 36 meses quando frases *jabberwocky* foram testadas vs. frases padrão das línguas.

Neste artigo, foi apresentada a técnica de ERP, mostrando o caminho que o campo tomou ao tentar entender os mecanismos subjacentes à cognição e especificamente à linguagem - a literatura adulta foi brevemente apresentada, mostrando seus principais componentes e estudos mais relevantes. Em seguida, apresentamos a literatura sobre o desenvolvimento infantil utilizando a técnica de ERP e seus desafios e resultados encorajadores. No final do artigo, os artigos mais importantes sobre a estrutura de argumentos usando a técnica ERP também foram resumidos. A ativação observada quando as crianças processam sentenças não ocorre necessariamente na mesma topografia e janela temporal encontradas em componentes clássicos de ERP. Alguns autores relataram ativações ocorrendo em janelas de 500-1500ms e em posições anteriores como P600. Isso estaria em consonância com o que Gouvea *et al.* (2010) observa em sua revisão da P600. No entanto, ainda não sabemos se todos os elementos relatados, como P600 em adultos, serão encontrados na população infantil e se eles representarão os mesmos processos subjacentes. É importante continuar a pesquisa para que possamos ter um corpo maior de dados de outros laboratórios, outras tarefas e línguas para saber mais sobre o processamento sintático das crianças.

Sendo assim, este campo de investigação tem grandes desafios a enfrentar ao tentar abranger diferentes fontes de informação de diferentes disciplinas relacionadas ao estudar a linguagem e/ou o desenvolvimento infantil. Mesmo assim, estou otimista de que a fertilização mútua (*cross-fertilization*) é possível (POEPPEL, EMBICK, 2005), se tentarmos pesquisar fenômenos com a devida diligência. No entanto, não devemos ser afoitos em supor que uma ativação cortical em crianças é diretamente comparável aos resultados encontrados em adultos. Até agora, ainda

precisamos de mais pesquisas para ter certeza de que as mesmas janelas de tempo e polaridade significam as mesmas computações para ambas populações.

Agradecimentos

Parte desta pesquisa foi conduzida com o apoio do CNPq e Fundação CAPES (Bolsista CAPES no Processo 17638-12-7), que apoiou um ano de trabalho em Paris, no Laboratoire de Science Cognitive et Psycholinguistique (LSCP) na École Normale Supérieure, Paris, França. Gostaria também de agradecer as minhas orientadoras Professoras Doutoras Aniela Imbrota França e Anne Christophe pelo suporte durante o período do meu doutorado e para além desse momento de formação acadêmica bem como toda a equipe dos laboratórios Acesin e LSCP.

Referências

- BERKUM, J.J.A.; HAGOORT, P.; BROWN, C. M. Semantic Integration in Sentences and Discourse: Evidence from the N400. *Journal of Cognitive Neuroscience*, MIT, v.11:6, p. 657-671, 1999. <https://doi.org/10.1162/089892999563724>.
- BERNAL, S.; DEHAENE-LAMBERTZ, G.; MILLOTTE, S.; CHRISTOPHE, A. Two-year-olds compute syntactic structure on-line. *Developmental Science*, Wiley Online Library, v. 13, n. 1, p. 69-76, jan. 2010. ISSN 1467-7687.
- BORNKESSEL-SCHLESEWSKY, I.; SCHLESEWSKY, M. An Alternative perspective on “semantic P600” effects in language comprehension, *Brain Research Reviews*, Elsevier, v.59, p. 55-73, 2008. <https://doi.org/10.1016/j.brainresrev.2008.05.003>.
- BROWN, C.; HAGOORT, P.; KUTAS, M. Postlexical Integration Processes in Language Comprehension: Evidence from Brain-Imaging Research. In: GAZZANIGA, M. S. (Ed.). *The New Cognitive Neurosciences*. [S.l.]: A Bradford Book; MIT Press, 2000. cap. 61, p. 881-895. ISBN 0-262-07195-9.

BRUSINI, P. *Découvrir les noms et les verbes: Quand les classes sémantiques initialisent les catégories syntaxiques*. 2012. Tese (Doutorado) – École Normale Supérieure, França, 2012.

BRUSINI, P.; DEHAENE-LAMBERTZ, G.; van HEUGTEN, M.; CARVALHO, A.; GOFFINET, F.; FIÉVET, A. C.; CHRISTOPHE, A. Ambiguous function words do not prevent 18-month-olds from building accurate syntactic category expectations: an ERP study. *Neuropsychologia*, Elsevier. In Press. <https://doi.org/10.1016/j.neuropsychologia.2016.08.015>.

BRUSINI, P.; DEHAENE-LAMBERTZ, G.; DUTAT, M.; GOFFINET, F.; CHRISTOPHE, A. ERP evidence for on-line syntactic computations in 2-year-olds. *Developmental Cognitive Neuroscience*, Elsevier, v. 19, p. 164-173, 2016a. <https://doi.org/10.1016/j.dcn.2016.02.009>.

BRUSINI, P. DEHAENE-LAMBERTZ, G.; DUTAT, M.; GOFFINET, F.; CHRISTOPHE, A. ERP evidence for on-line syntactic computations in 2-year-olds, *Developmental Cognitive Neuroscience*, Elsevier, v.19, p. 164-173, jun. 2016b. <https://doi.org/10.1016/j.dcn.2016.02.009>.

COCH, D.; MITRA, P.; GEORGE, E. Behavioral and ERP evidence of word and pseudoword superiority effects in 7- and 11-year-olds. *Brain Research*, Elsevier, v. 1486, p. 68-81, nov. 2012. ISSN 1872-6240.

CONBOY, B. T.; RIVERA-GAXIOLA, M.; SILVA-PEREYRA, J. F.; AND KUHL, P. K. Event-related potential studies of early language processing at the phoneme, word, and sentence levels. In: FRIEDERICI, A. D.; THIERRY, G. (Org.). *Early language development: bridging brain and behavior*, Trends in language acquisition research series. Amsterdam/The Netherlands: John Benjamins, 2008. v. 5, p. 24-64. <https://doi.org/10.1075/tilar.5.04con>.

COSTA, M. U. C. L. M. *Argument Structure in Language Acquisition: an ERP Study*. 2015. 174f. Tese (Doutorado em Linguística) –Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015.

COULSON, S.; KING, J. W.; KUTAS, M. Expect the Unexpected: Event-related Brain Response to Morphosyntactic Violations. *Language and Cognitive Processes*, Taylor & Francis Online, v. 13, n. 1, p. 21-58, 1998. ISSN 0169-0965.

DEBOER, T.; SCOTT, L.; NELSON, C. Methods for acquiring and analyzing infant event-related potentials. In: HAAN, M. D. (Org.). *Infant EEG and event-related potentials*. [S.l.]: Psychology Press, 2007. cap. 1, p. 5-38.

DEBRUILLE, J. B. The N400 potential could index a semantic inhibition. *Brain Research Reviews*, Elsevier, v. 56, n. 2, p. 472-427, dez. 2007. ISSN 0165-0173.

DEHAENE-LAMBERTZ, G.; PENA, M. Electrophysiological evidence for automatic phonetic processing in neonates. *Neuroreport*, LWW Journals, v. 12, n. 14, p. 3155-3158, out. 2001. ISSN 0959-4965.

FRANCA, A. I. *et al.* Discriminating among different types of verb-complement merge in Brazilian Portuguese: An ERP study of morpho-syntactic sub-processes. *Journal of Neurolinguistics*, Elsevier, v. 17, p. 425-437, 2004. ISSN 09116044.

FRANÇA, A. I. *Concatenações linguísticas: estudo de diferentes módulos cognitivos na aquisição e no córtex*. 236f. 2002. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2002.

FRANÇA, A; GOMES, J. Técnica de ERP: investigando a assimetria sujeito-objeto na interface sintaxe-semântica com EEG. *Letras de Hoje*, Porto Alegre, v. 50, n. 3, p. 360-370, jul.-set. 2015. <https://doi.org/10.15448/1984-7726.2015.3.18411>.

FRANÇA, A.I.; LEMLE, M; GESUALDI, A.; CAGY, M.; INFANTOSI, A.F.C. A neurofisiologia do acesso lexical: palavras em português. *Revista Veredas*, Juiz de Fora, UFJF, v. 2, p. 34-49, 2008.

FRIEDERICI, A. D. The time course of syntactic activation during language processing: a model based on neuropsychological and neurophysiological data. *Brain and Language*, Elsevier, v. 50, p. 259-281, 1995. <https://doi.org/10.1006/brln.1995.1048>

FRIEDERICI, A. D. Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, Elsevier, v. 6, n. 2, p. 78-84, fev. 2002. ISSN 1364-6613.

FRIEDERICI, A. D. *et al.* Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebral Cortex* [New York, N.Y., 1991], Oxford Academic, v. 16, n. 12, p. 1709-1717, dez. 2006. ISSN 1047-3211.

FRIEDERICI, A. D.; FRIEDRICH, M.; CHRISTOPHE, A. Brain responses in 4-month-old infants are already language specific. *Current Biology: CB*, Bethesda, v. 17, n. 14, p. 1208-1211, jul. 2007. ISSN 0960-9822.

FRIEDERICI, A. D.; FRISCH, S. Verb Argument Structure Processing: The Role of Verb Specific and Argument-Specific Information. *Journal of Memory and Language*, Elsevier, v. 43, n. 3, p. 476-507, out. 2000. ISSN 0749596X.

FRIEDERICI, A. D. *et al.* Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research*, Bethesda, v. 11, p. 305-323, 2001. ISSN 09266410.

FRIEDERICI, A. D.; MEYER, M. The brain knows the difference: two types of grammatical violations. *Brain Research*, v. 1000, n. 1-2, p. 72-77, mar. 2004. ISSN 0006-8993.

FRIEDERICI, A. D.; PFEIFER, E.; HAHNE, A. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, Bethesda, v. 1, n. 3, p. 183-192, out. 1993. ISSN 09266410.

FRIEDERICI, A. D.; THIERRY, G. (Org.). *Early Language Development. Bridging brain and behaviour.* (Trends in Language Acquisition Research) Amsterdam/Philadelphia, John Benjamins Publishing Company, 2008. <https://doi.org/10.1075/tilar.5>.

FRIEDERICI, A. D.; WEISSENBORN, J. Mapping sentence form onto meaning: the syntax-semantic interface. *Brain research*, Elsevier, v. 1146, p. 50-58, maio 2007. ISSN 0006-8993.

FRIEDRICH, M.; FRIEDERICI, A. N400-like semantic incongruity effect in 19-month-old infants: Processing known words in picture contexts. *Journal of cognitive neuroscience*, p. 1465-1477, 2004. <https://doi.org/10.1162/0898929042304705>.

FRIEDRICH, M.; FRIEDERICI, A. D. Maturing brain mechanisms and developing behavioral language skills. *Brain and Language*, Elsevier, v. 114, n. 2, p. 66-71, 2010. ISSN 0093934X.

- FRISCH, S.; HAHNE, A.; FRIEDERICI, A. D. Word category and verb-argument structure information in the dynamics of parsing. *Cognition*, Elsevier, v. 91, n. 3, p. 191-219, abr.2004. ISSN 0010-0277.
- GESUALDI, A. DA R.; FRANÇA, A. I. Event-related brain potentials (ERP): an overview. *Revista Linguística, UFRJ, Rio de Janeiro*, v. 7, n. 2, p. 24-42, 2011. 39.
- GIBSON, E. Linguistic complexity: locality of syntactic dependencies. *Cognition*, Elsevier, v. 68, p. 1-76, 1998. ISSN 00100277.
- GOMES, J. N. *Investigating the dissociation between n400 and p600 effects on the syntax semantics interface: an erp study*. 183 f. Tese (Doutorado) – Faculdade de Letras da UFRJ, Rio de Janeiro, 2014.
- GOMES, J. N.; FRANÇA, A. I. The ample semantic scope of minute language computations: an ERP study of words in Portuguese. *Revista da ANPOLL*, Anpoll, v. 34, p. 309-346, 2013. Disponível em: <<http://anpoll.emnuvens.com.br/revista/article/view/673/698>> 45>.
- GOMES, J.; FRANÇA, A. I. A técnica de ERP: Investigando a assimetria sujeito-objeto na interface sintaxe-semântica com EEG. *Letras de Hoje* (Online), PUCRS, Porto Alegre, v. 4, p. 152-163, 2015.
- GOUVEA, A. C. *et al.* The linguistic processes underlying the P600. *Language and Cognitive Processes*, Taylor & Francis Online, v. 25, n. 2, p. 149-188, fev. 2010. ISSN 0169-0965.
- HAGOORT, P. How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. *Neuroimage*, Elsevier, v. 20, p. 18-29, 2003. <https://doi.org/10.1016/j.neuroimage.2003.09.013>.
- HAGOORT, P. On Broca, brain, and binding: a new framework. *Trends in cognitive sciences*, Elsevier, v. 9, n. 9, p. 416-423, set. 2005. ISSN 1364-6613.
- HAGOORT, P.; BROWN, C.; GROOTHUSEN, J. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, Taylor & Francis Online, v. 8, n. 912873514, p. 439-483, 1993. ISSN 0169-0965.
- HAGOORT, P.; BROWN, C.; OSTERHOUT, L. The neurocognition of syntactic processing. In: BROWN, C. M.; HAGOORT, P. (Org.). *The neurocognition of language*. Oxford: Oxford University Press, 1999. p.273-317.

HAGOORT, P.; BROWN, C. M. ERP effects of listening to speech compared to reading: the P600 / SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia*, Elsevier, v. 38, p. 1531-1549, 2000. [https://doi.org/10.1016/S0028-3932\(00\)00053-1](https://doi.org/10.1016/S0028-3932(00)00053-1).

HAGOORT, P.; LEVELT, W. The speaking brain. *Science*, American Association for the Advancement of Science, v. 326, p. 372-373, 2009. <https://doi.org/10.1126/science.1181675>.

HÄNEL-FAULHABER, B.; SKOTARA, N.; KÜGOW, M.; SALDEN, U.; BOTTARI, D.; RÖDER, B. ERP correlates of German Sign Language processing in deaf native signers. *BMC Neuroscience*, Springer Nature, v.15, n. 62, p. 1-11, 2014.

HERTEN, M. van; KOLK, H. H. J.; CHWILLA, D. J. An ERP study of P600 effects elicited by semantic anomalies. *Brain research. Cognitive Brain Research*, Bethesda, v. 22, n. 2, p. 241-255, mar. 2005. ISSN 0926-6410.

HOEKS, J. C. J.; STOWE, L. a.; DOEDENS, G. Seeing words in context: the interaction of lexical and sentence level information during reading. *Brain Research. Cognitive Brain Research*, Bethesda, v. 19, n. 1, p. 59-73, mar. 2004. ISSN 0926-6410.

HOLCOMB, P.; COFFEY, S.; NEVILLE, H. Visual and auditory sentence processing: A developmental analysis using event-related brain potentials. *Developmental Neuropsychology*, Taylor & Francis Online, v. 8, p. 203-312, 1992. <https://doi.org/10.1080/87565649209540525>.

HWANG, H.; STEINHAEUER, K. Phrase length matters: the interplay between implicit prosody and syntax in Korean “garden path” sentences. *Journal of Cognitive Neuroscience*, MIT Press, v. 23, n. 11, p. 3555-3575, nov. 2011. ISSN 1530-8898.

JOHNSON, M. Functional Brain Development During Infancy. In: BREMMER, G.; FOGEL, A. (Ed.). *Blackwell handbook of infant development*. [S.l.]: Blackwell Publishing, 2001. cap. 6.

KAAN, E. *et al.* The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, Taylor & Francis Online, v. 15, n. 2, p. 159-201, abr. 2000. ISSN 0169-0965.

KIM, A.; OSTERHOUT, L. The independence of combinatorial semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, Elsevier, v. 52, n. 2, p. 205-225, fev. 2005. ISSN 0749596X.

KOLK, H. H. J. *et al.* Structure and limited capacity in verbal working memory: a study with event-related potentials. *Brain and language*, Elsevier, v. 85, p. 1-36, 2003.

KOOIJMAN, V. *et al.* Predictive brain signals of linguistic development. *Frontiers in Psychology*, Frontiers Media, v. 4, p. 25, Feb. 2013. ISSN 1664-1078. <https://doi.org/10.3389/fpsyg.2013.00025>.

KOTZ, S. A.; HOLCOMB, P. J.; OSTERHOUT, L. ERPs reveal comparable syntactic sentence processing in native and non-native readers of English. *Acta Psychologica*, Elsevier, v. 128, n. 3, p. 514-527, jul. 2008. ISSN 0001-6918.

KOVELMAN, I. Neuroimaging Methods. In: HOFF, E. *Research Methods in Child Language: A Practical Guide*. [S.l.]: Wiley-Blackwell, 2012. Cap. 4, p. 44-59.

KUPERBERG, G.; SITNIKOVA, T. Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research*, Bethesda, v. 17, p. 117-129, 2003. [https://doi.org/10.1016/S0926-6410\(03\)00086-7](https://doi.org/10.1016/S0926-6410(03)00086-7).

KUPERBERG, G. R. Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, Elsevier, v. 1146, p. 23-49, maio 2007. ISSN 0006-8993.

KUPERBERG, G. R. *et al.* Building up linguistic context in schizophrenia: evidence from self-paced reading. *Neuropsychology*, Elsevier, v. 20, n. 4, p. 442-452, jul. 2006. ISSN 0894-4105.

KUTAS, M.; DELONG, K. A Sampler of Event-related brain Potential (ERP) Analyses of Language Processing. *Brain Research in Language*, Springer, v. 1, p. 153-186, 2008. https://doi.org/10.1007/978-0-387-74980-8_6.

KUTAS, M.; FEDERMEIER, K. D. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, Annual Reviews, v. 62, p. 621-647, jan. 2011. ISSN 1545-2085.

KUTAS, M.; HILLYARD, S. Reading senseless sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, American Association for the Advancement of Science, v. 207, n. 4427, p. 203-205, jan. 1980. <https://doi.org/10.1126/science.7350657>.

LAGE, A. C. Aspectos neurofisiológicos de concatenação e idiomaticidade em português do Brasil: um estudo de potenciais bioelétricos relacionados a eventos lingüísticos (ERPs). 2005. Tese (Doutorado em Lingüística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

LAU, E. F.; PHILLIPS, C.; POEPPPEL, D. A cortical network for semantics: (de)constructing the N400. *Nature Reviews. Neuroscience*, Macmillan Publishers, v. 9, n. 12, p. 920-933, dez. 2008. ISSN 1471-0048.

LUCK, S. J. *An Introduction to the Event-Related Potential Technique*. [S.l.]: MIT Press, 2005. ISSN 00393665.

MARANTZ, A. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review*, De Gruyter, n. 22, p. 429-445, Jan. 2005. ISSN 0167-6318.

MILLS, D. L. *et al.* Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience*, MIT Press, v. 16, n. 8, p. 1452-1464, out. 2004. ISSN 0898-929X.

NELSON, C.; LUCIANA, M. *Handbook of Developmental Cognitive Neuroscience*. 2nd. ed. [S.l.]: MIT Press, 2008. ISBN 9780262141048.

NEVINS, A. *et al.* The role of feature-number and feature-type in processing Hindi verb agreement violations. *Brain Research*, Elsevier, v. 1164, p. 81-94, 2007. <https://doi.org/10.1016/j.brainres.2007.05.058>.

OBERECKER, R.; FRIEDERICI, A. D. Syntactic event-related potential components in 24-month-olds' sentence comprehension. *Neuroreport*, LWW Journals, v. 17, n. 10, p. 1017-1021, jul. 2006. ISSN 0959-4965.

OBERECKER, R.; FRIEDRICH, M.; FRIEDERICI, A. D. Neural correlates of syntactic processing in two-year-olds. *Journal of Cognitive Neuroscience*, MIT Press, v. 17, n. 10, p. 1667-1678, out. 2005. ISSN 0898-929X.

OSTERHOUT, L.; HOLCOMB, P. J. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, Elsevier, v. 31, n. 6, p. 785-806, dez. 1992. ISSN 0749596X.

OSTERHOUT, L.; HOLCOMB, P. J.; SWINNEY, D. a. Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, Washington, v. 20, n. 4, p. 786-803, jul. 1994. ISSN 0278-7393.

OSTERHOUT, L.; KIM, A.; KUPERBERG, G. The neurobiology of sentence comprehension. In: SPIVEY, M.; JOANNISSE, M.; MCRAE, K. (Org.). *The Cambridge Handbook of Psycholinguistics*. Cambridge: Cambridge University Press, 2007. p. 1-23.

OSTERHOUT, L.; MCLAUGHLIN, J. Novice Learners, Longitudinal Designs, and Event-Related Potentials: A Means for Exploring the Neurocognition of Second Language Processing. *Language Learning*, Wiley Online Library, v. 56, Supplement 1, p. 4-25, 2006.

PHILLIPS, C.; KAZANINA, N.; ABADA, S. H. ERP effects of the processing of syntactic long distance dependencies. *Brain research. Cognitive Brain Research*, Bethesda, v. 22, n. 3, p. 407-28, mar. 2005. ISSN 0926-6410.

PICTON, T. W. *et al.* Guidelines for using human event-related potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, Wiley Online Library, v. 37, n. 2, p. 127-152, mar. 2000. ISSN 0048-5772.

POEPEL, D.; EMBICK, D. Defining the relation between linguistics and neuroscience. In: CUTLER, A. (Ed.). *Twenty-first century psycholinguistics: four cornerstones*. Mahwah, N.J.: Lawrence Erlbaum, 2005.

POEPEL, D.; OMAKI, A. Language acquisition and ERP approaches: Prospects and challenges. In: FRIEDERICI, A.; THIERRY, G. *Early language development: bridging brain and behavior*. Amsterdam: John Benjamins, 2008. <https://doi.org/10.1075/tilar.5.12poe>.

PUJOL, J. *et al.* Myelination of language-related areas in the developing brain. *Neurology*, American Academy of Neurology, v. 66, n. 3, p. 339-343, fev. 2006. ISSN 1526-632X.

PYLKKÄNEN, L.; BRENNAN, J.; BEMIS, D. Grounding the cognitive neuroscience of semantics in linguistic theory. *Language and Cognitive Processes*, Taylor & Francis Online, v. 26, n. 9, p. 1317-1337, 2011.

SHAFER, V.; GARRIDO-NAG, K. The neurodevelopmental bases of language. In: HOFF, E.; SHATZ, M. (Org.). *Blackwell handbook of language Development*. Oxford: Blackwel, 2007. cap. 2, p. 21-45. <https://doi.org/10.1002/9780470757833.ch2>.

SILVA-PEREYRA, J. *et al.* Grammatical processing without semantics? An event related brain potential study of preschoolers using jabberwocky sentences. *Journal of Cognitive Neuroscience*, MIT Press, v. 19, n. 6, p. 1050-1065, jun. 2007. ISSN 0898-929X.

SILVA-PEREYRA, J. F. *et al.* Sentence processing in 30-month-old children: an event-related potential study. *Neuroreport*, LWW Journals, v. 16, n. 6, p. 645-648, abr. 2005a. ISSN 0959-4965.

SILVA-PEREYRA, J. F.; RIVERA-GAXIOLA, M.; KUHL, P. K. An event-related brain potential study of sentence comprehension in preschoolers: semantic and morphosyntactic processing. *Brain Research. Cognitive Brain Research*, Elsevier, v. 23, n. 2-3, p.247-258, maio 2005b. ISSN 0926-6410.

TEPLAN, M. Fundamentals of EEG measurement. *Measurement Science Review*, De Gruyter, v. 2, p. 1-11, 2002.

ZHANG, Y. *et al.* Syntax does not necessarily precede semantics in sentence processing: ERP evidence from Chinese. *Brain and Language*, Elsevier, v. 126, n. 1, p. 8-19, jul. 2013. ISSN 1090-2155.

YUAN, S.; FISHER, C. Really? She Blicked the Baby? Two-Year-Olds Learn Combinatorial Facts About Verbs. *Psychological Science*, Sage Journals, v. 20, n. 5, p. 619-626, 2009.

**Aquisição de estruturas possessivas inalienáveis:
o caso dos nomes de parte do corpo em inglês americano
e português brasileiro**

***Acquisition of inalienable possessive structures: the case of body-
part names in American English and Brazilian Portuguese***

Fernanda Mendes

Universidade Estadual de Campinas, Campinas, São Paulo / Brasil

fernanda.mds@gmail.com

Resumo: Este estudo testa a aquisição de estruturas de posse inalienável em português brasileiro e inglês americano, assumindo que crianças falantes nativas de inglês americano teriam disponível, num primeiro momento, a mesma gramática disponível para crianças falantes nativas de português brasileiro, em que a leitura inalienável poderia ser vinculada tanto na presença de pronomes possessivos quanto na presença de determinantes definidos com nomes de parte do corpo. Aplicando a Tarefa de Julgamento de Valor de Verdade observou-se que a distinção entre as gramáticas das duas línguas apareceu por volta dos seis anos de idade, quando crianças falantes de inglês americano restringiram o uso do determinante definido de acordo com a gramática alvo, não permitindo o seu uso nas estruturas contendo nomes de partes do corpo, quando foi veiculada a interpretação inalienável. Neste caso, o determinante definido foi substituído pelo pronome possessivo.

Palavras-chave: posse; inalienável; aquisição.

Abstract: The current study tests the acquisition of inalienable possessive structures in Brazilian Portuguese and American English, assuming that American English-speaking children start with the same grammar as

Brazilian Portuguese-speaking children do. This first grammar would allow inalienable interpretation being carried by possessive pronouns as well as definite determiners introducing body-parts names. Using Truth Value Judgement Task method it is noticed that the difference between Brazilian Portuguese grammar and American English grammar appeared around the age of six, when American English-speaking children approach the target grammar, in which the presence of the possessive pronoun is what allows the inalienable interpretation and the use of definite determiners is restricted to alienable possession constructions with body-parts names.

Keywords: possession; inalienable; acquisition.

Recebido em: 10 de dezembro de 2016

Aprovado em: 19 de abril de 2017

1 Introdução

De forma geral, as construções de posse inalienável são definidas como estruturas possessivas especiais em que o possuído e o possuidor estabelecem uma relação de posse que não pode ser transferida a outro possuidor qualquer. Usualmente, essas estruturas contam com nomes de partes do corpo, tais como *pernas* ou *nariz*, ou relacionais, como *pai* e *irmãos*,¹ e implementam uma relação possessiva, independentemente da presença de um pronome que a expresse – tal como em *O Pedro lavou a mão (dele)* ou *O Pedro abraçou a mãe (dele)*.²

A fim de investigar como esse tipo de construção possessiva se comporta no PB adulto e se desenvolve na gramática infantil, este trabalho se apoia em estudos realizados anteriormente que examinam como esse mesmo tipo de estrutura funciona em outras línguas – não

¹ No entanto, apenas os nomes de partes do corpo serão examinados no presente artigo.

² É importante observar que estas sentenças, em PB, também podem veicular a interpretação alienável. Ou seja, é possível haver uma leitura em que o objeto possuído *a mão (dele)* ou *a mãe (dele)* não pertença ao possuidor disponível na sentença *O Pedro*. No entanto, esta não é a interpretação requerida na presente discussão sobre inalienabilidade.

necessariamente pertencentes à mesma família –, tais como o francês, o espanhol e o inglês.

Assim, serão apresentadas, ao longo deste artigo, pesquisas que tratam sincronicamente da posse inalienável e da sua aquisição, além de estudos que examinam fenômenos relacionados a ela – tais como a aquisição de determinantes e de Princípios A e B da Teoria da Ligação (CHOMSKY, 1981).

De acordo com esses trabalhos, serão apresentadas as hipóteses e previsões aventadas no presente estudo, a serem avaliadas pelos experimentos aplicados em inglês americano (AE) e português brasileiro (PB), exibidos em seguida.

Na seção 2 a seguir, é apresentado o viés teórico deste artigo, incluindo o que a literatura da área afirma sobre a posse inalienável na gramática adulta e infantil, bem como as hipóteses e previsões da presente pesquisa. Na seção 3, em seguida, é apresentado o viés experimental deste artigo, incluindo a metodologia adotada, os fatores envolvidos e os testes específicos (gerados a partir da combinação desses fatores), os sujeitos participantes, os resultados e a discussão. E, por fim, na seção 4 é traçada a conclusão deste artigo.

2 Estruturas possessivas inalienáveis: viés teórico

2.1 PB adulto: estratégias sincrônicas para a marcação de posse inalienável

Segundo pesquisas diacrônicas sobre posse inalienável³, o pronome clítico dativo de posse não se constitui mais como uma estratégia sintática possível para veicular a interpretação inalienável da estrutura em PB. Atualmente, são empregadas outras estratégias, como o uso do pronome possessivo e do determinante definido, que serão analisadas nesta seção, com foco especial na última, já que essa estratégia é a que parece diferenciar o PB de outras línguas românicas, como o francês e o espanhol, além de o distinguir do AE.

Segundo Guéron (1985), a interpretação inalienável é veiculada quando o sintagma nominal que denota a coisa possuída é uma parte

³ Ver MENDES, 2012.

intrínseca do sintagma nominal possuidor, que pode ser o sujeito, como em (1),⁴ ou o dativo possuidor, como (2).⁵

- (1) a. *Jean lève la main.*
 b. *O João levanta a mão.*
- (2) a. *Je lui ai coupé les cheveux.*
 b. *Eu lhe cortei os cabelos.*

Por outro lado, a interpretação alienável é veiculada quando o sintagma nominal que denota a coisa possuída é “disjunto” do sujeito ou do dativo possuidor, como mostram os exemplos do AE em (3).⁶

- (3) a. *#John raises the hand.*
 b. *#I cut the hair for her.*

Segundo a classificação proposta pela autora, construções em que o sintagma nominal contendo um nome de parte do corpo e introduzido por um determinante definido está em posição do objeto do verbo, como em (1)-(3), apresentam ambiguidade entre a leitura alienável e a inalienável em línguas como o francês, enquanto, em línguas como o AE, há apenas a leitura alienável disponível. Conforme os exemplos (1b) e (2b), observa-se que os dados do PB se comportam tais quais os do francês com respeito a essa classificação.

No caso do AE, para que a leitura inalienável esteja disponível é preciso que, no lugar do determinante definido, haja um pronome possessivo,⁷ tal como ilustra (4a).

- (4) a. *John raises his hand.*
 b. *Jean lève sa main.*
 c. *João levanta a mão dele.*

⁴ Exemplo retirado de GUÉRON, 1985, p. 43.

⁵ Exemplo retirado de GUÉRON, 1985, p. 43.

⁶ Exemplos retirados de GUÉRON, 1985, p. 43-44.

⁷ Para uma análise transformacional dessa estrutura, ver CHOMSKY, 1970, p.200.

Os exemplos em (4b-c) mostram que essa possibilidade também está disponível no francês e no PB. Portanto, quando a estrutura envolve um pronome possessivo, ambas as leituras, alienável e inalienável, estão disponíveis, já que, em (4), *his hand*, *sa main* e *a mão dele* podem tanto pertencer ao sujeito da sentença – *John/Jean/João* –, quanto a um outro possuidor qualquer, que pode ter sido mencionado no discurso.

Assim, o Quadro 1 apresenta as leituras veiculadas por cada tipo de determinante que introduz o sintagma nominal contendo o nome possuído em cada língua observada:

QUADRO 1 – Leitura permitida com o determinante: AE e francês/PB

	Determinante definido	Pronome possessivo
Leitura inalienável	PB/Francês	PB/Francês e AE
Leitura alienável	PB/Francês e AE	PB/Francês e AE

Segundo a hipótese da autora, a diferença apresentada, contrapondo os dados do AE aos dados do francês ou do PB, em relação à disponibilidade das leituras inalienável e alienável nos exemplos introduzidos por determinantes definidos seria consequência do valor atribuído a um parâmetro morfofonológico, chamado pela autora de “parâmetro de inclusão de PRO”, segundo o qual o determinante definido que constitui o sintagma inalienável pode ser uma anáfora pronominal ou um operador iota.

Como uma anáfora pronominal, ou seja, um pronome sem referência independente, o determinante definido seria responsável pela leitura inalienável vinculada, tal qual ocorre em línguas como o francês ou o PB.

Como um operador iota, ou seja, um operador semântico que representa a parte que denota a unicidade de uma descrição definida – isto é, a sua *definitude* –, o determinante definido seria responsável pela leitura alienável veiculada, tal qual acontece em línguas como o francês, o PB e o AE.

Evidência para esta hipótese, na aquisição, seriam as marcas de concordância (gênero, número e pessoa⁸) do determinante. Isto é, se a

⁸ Segundo Guéron (1985), o determinante definido, nas línguas românicas, tem o traço de 3ª pessoa.

língua tem os determinantes morfológicamente marcados, como em (5), o determinante definido pode funcionar, além de como um operador iota, como uma anáfora pronominal.

- | | | | |
|-----|----|-------------------|--------------|
| (5) | a. | <i>o(s)</i> | garoto(s) |
| | | [3Sg. M.Sg.(Pl.)] | [M.Sg.(Pl.)] |
| | b. | <i>a(s)</i> | garota(s) |
| | | [3Sg. F.Sg.(Pl.)] | [F.Sg.(Pl.)] |
| (6) | a. | <i>the</i> | boy(s) |
| | | [Ø] | [M.Sg.(Pl.)] |
| | b. | <i>the</i> | girl(s) |
| | | [Ø] | [F.Sg.(Pl.)] |

Por outro lado, se a língua não tem os determinantes morfológicamente marcados, como em (6), o determinante definido não pode atuar como uma anáfora pronominal, mas apenas como um operador iota.

Sendo o determinante definido uma anáfora, Guéron (1985) propõe, como uma das condições sintáticas básicas, que ele se sujeite à ligação anafórica. Ou seja, para que haja a interpretação inalienável é preciso que estas construções estejam submetidas ao Princípio A da Teoria da Ligação, apresentado em (7), conforme exemplo (12A) em Chomsky (1981, p.188).⁹

- (7) Princípio A:
 “[...] uma anáfora tem que estar ligada em seu domínio de ligação”.

De acordo com esse princípio, a anáfora, precisa ter um antecedente que a c-comande e ambos precisam estar contidos no domínio

⁹ Do original: “An anaphor is bound in its governing category” (p. 188).

de ligação da primeira, sendo domínio de ligação definido como segue em (8), conforme Roberts (1997, p.142, exemplo 21').¹⁰

- (8) O domínio de ligação de α é o XP mínimo que contém α e
- a. Um sujeito que é distinto de α e que não contém α ; ou
 - b. A flexão que atribui Caso Nominativo para α

Assim, no exemplo em (1) acima, *la main/a mão* é uma anáfora e seu antecedente, o sintagma nominal possuidor *Jean/O João*, além de c-comandar a anáfora, se configura como um sujeito distinto dela que não a contém. Sendo assim, é explicada a impossibilidade da vinculação da leitura inalienável em exemplos como (9) abaixo, já que o Princípio A não seria respeitado.

- (9) *O dente caiu.*

Além disso, é necessário que haja uma cadeia lexical, que Guéron (1985), de acordo com Chomsky (1981), define como:

- (10) “[...] um conjunto de dois ou mais nominais relacionados por ligação anafórica e interpretados como um único argumento em LF” (GUÉRON, 1985, p. 44).

De acordo com a autora, todos os tipos de cadeia são afetados pela “Restrição de não-distintividade” enunciada em (11) abaixo, o que a torna uma condição a ser respeitada também para o estabelecimento da interpretação inalienável.

- (11) “Se A e B são elos de uma cadeia, então o referente de A é não-distinto do referente de B” (GUÉRON, 1985, p.44).

A junção da noção de cadeia lexical apresentada em (10) com a restrição de não-distintividade enunciada em (11) resulta, segundo

¹⁰ Do original: “*Binding Domain (Revised)* The binding domain (BD) of α is the smallest XP containing α and:

- EITHER (a) a subject (distinct from α and which does not contain α)
OR (b) the I that assigns Nominative Case to α ” (p. 142).

a autora, em uma simplificação da gramática frente aos modelos propostos anteriormente, fazendo com que nada mais seja necessário gramaticalmente para que se estabeleça uma interpretação inalienável.

Vergnaud e Zubizarreta (1992), no entanto, discordam desta hipótese, já que uma análise baseada na teoria de controle não capturaria certas restrições que se aplicam a essas estruturas. Portanto, eles propõem que a diferença apresentada entre a disponibilidade das leituras alienável e inalienável em AE e francês seja consequência da dupla interpretação que o sintagma nominal definido contendo nomes de partes do corpo, por exemplo, pode veicular. Segundo estes autores, sintagmas nominais definidos e referenciais podem expressar interpretação de *token* e interpretação de *type* de acordo com a “Lei de correspondência”, apresentada em (12) abaixo.

(12) Lei de correspondência

“Quando um DP ou um NP denota, o DP denota um *token* e o NP denota um *type*”¹¹.

A interpretação de *token*, que licencia a interpretação alienável, corresponde a um DP que tem por núcleo um *determinante substantivo*, conforme Longobardi (1994), resultando em um DP “completo”. Já a interpretação de *type*, que licencia a interpretação inalienável, corresponde tanto a um NP nu, quanto a um DP que tem por núcleo um *determinante expletivo*.

Os autores definem o determinante expletivo como “uma categoria que não tem representação em um certo domínio D”¹² (VERGNAUD e ZUBIZARRETA, 1992, p. 595), conforme a definição de “domínio D” adotada em Chomsky (1981, p.324).

Longobardi (1994), propondo um princípio para reger a realização do núcleo da projeção do determinante, apresentado em (13) abaixo, afirma que determinantes definidos podem ser divididos entre determinantes substantivos e determinantes expletivos, podendo ser a última categoria ainda subdividida entre aqueles que introduzem nomes genéricos massivos plurais e aqueles que introduzem nomes próprios.

¹¹ Do original: “*Correspondence Law* when a DP or an NP denotes, the DP denotes a token and the NP denotes a type” (p. 612, exemplo 45).

¹² Do original: “an expletive in that sense [Chomsky, 1981] is a category that has no representation in domain D” (p. 595).

- (13) “A realização fonológica da posição D é licenciada apenas se ela expressa conteúdo semântico ou traços gramaticais, ou como último recurso”¹³ (p. 654, exemplo 89).

Segundo Castro (2006, 2007), observa-se que, em PB, além de o determinante expletivo, nulo ou lexicalizado¹⁴, ocorrer introduzindo sintagmas nominais com nomes próprios e nomes genéricos, esses determinantes também podem introduzir sintagmas nominais possessivos¹⁵ (no caso, sintagmas nominais introduzidos por possessivos simples prenominais), como mostram os exemplos em (14) abaixo.

- (14) a. (O) *João* comprou um livro.
 b. (As) *crianças* gostam de gibis.
 c. (O) *meu* livro foi publicado.

Assim, a interpretação de *type* estaria disponível para línguas como o PB, que apresentam determinantes expletivos, e não para o AE, na qual este tipo de determinantes não estão presentes¹⁶, como evidenciam os exemplos em (15)¹⁷ abaixo.

¹³ Do original: “The phonetic realization of the D position is licensed only if it expresses semantic content or grammatical features, or as a last resort” (p. 654, exemplo 89).

¹⁴ Segundo Castro (2006, 2007), no PB o determinante expletivo pode ocorrer tanto nulo quanto lexicalizado, enquanto, no PE, ele ocorre apenas lexicalizado.

¹⁵ Para uma análise diacrônica dessa coocorrência, ver os estudos de Vianna (2004), Floripi (2008) e Rinke (2010).

¹⁶ Segundo Longobardi (1994), há dois casos em AE em que o determinante definido expletivo pode ocorrer: introduzindo sintagmas nominais com nomes genéricos não-massivos e singulares, como em (i) abaixo (LONGOBARDI, 1994, p. 650, exemplo 80a-b); ou introduzindo sintagmas nominais com adjetivos substantivados genéricos, como em (ii) abaixo (LONGOBARDI, 1994, p.644, exemplo 44a-b).

- (i) a. *The lion* has four legs.
 b. **Lion* has four legs.
 (ii) a. *The rich* are becoming even richer.
 b. **Rich* are becoming even richer.

¹⁷ Exemplos retirados de Longobardi (1994, p. 631, exemplo 43), exceto exemplo 39d.

- (15) a. *I love *the France*.
 b. **The beavers* are mammals.
 c. **The wine* is made out of grapes.
 d. **The my* book was published.

A interpretação de *token*, por outro lado, estaria disponível para ambos os tipos de línguas, dado que o PB e o AE apresentariam igualmente determinantes substantivos.

Dessa forma, a gramática adulta do AE e do PB apresentariam padrões opostos em relação à leitura permitida quando está presente o determinante definido: enquanto, na primeira, determinantes definidos, por serem substantivos, seriam usados unicamente para denotar a leitura alienável, na segunda, esses mesmos determinantes seriam usados para denotar ambas as leituras, sendo inalienável quando o determinante definido é expletivo e alienável quando o determinante definido é substantivo, como ilustra o Quadro 2 abaixo.

QUADRO 2 – Leitura permitida com o determinante definido: AE e PB

Língua	Determinante substantivo	Determinante expletivo
PB	Leitura alienável	Leitura inalienável
AE	Leitura alienável	<i>Não se aplica</i>

Já o uso dos pronomes possessivos, por outro lado, desencadearia, em ambas as línguas, as leituras alienável e inalienável, sendo, porém, em AE, a única opção no caso da última.

Vergnaud e Zubizarreta (1992) se apoiam ainda na dependência semântica. Segundo essa noção, o objeto inalienável é uma entidade dependente, já que ele é inerentemente definido em termos de outro objeto do qual faz parte. Contrariamente, o objeto alienável é uma entidade independente, isto é, tem definição própria.

Dessa forma, os autores assumem que um nome inalienável, diferentemente de um nome alienável, tome um argumento possuidor, que pode ser lexicalizado dentro ou fora do sintagma nominal. Logo, nomes de partes do corpo, por exemplo, teriam duas entradas lexicais:

- (16) a. mão (x)
b. mão

Uma em que ele toma um argumento possuidor e veicula a leitura inalienável, como em (16a), e outra em que ele não toma um argumento possuidor e veicula a leitura alienável, como em (16b) (Cf. AUTHIER, 1988; TELLIER, 1988 *apud* VERGNAUD; ZUBIZARRETA, 1992).

2.2 PB infantil: posse inalienável e determinantes definidos

Baseando-se nas análises exibidas na seção anterior para a gramática adulta do PB, nesta seção, serão apresentados alguns estudos, dentro dessa mesma língua, ou ainda examinando outras línguas, tais como o holandês, o espanhol e o AE, que tratam da gramática infantil focando na aquisição da posse inalienável juntamente com outros tipos de estruturas relacionadas diretamente a ela, como determinantes definidos.

Ramos (1999 *apud* PÉREZ-LEROUX *et al.*, 2002a, b), examinando a interpretação de determinantes definidos em contextos de posse inalienável na gramática de crianças com DEL,¹⁸ observa, de acordo com os resultados das crianças típicas falantes de AE, uma taxa decrescente na veiculação da leitura inalienável em estruturas em que está presente o determinante definido. Enquanto é encontrada uma taxa de 30% nos resultados do grupo mais jovem – 3;8 a 4;5 anos –, no grupo mais velho – 4;7 a 5;7 anos –, é encontrada uma taxa de 23%.

Resultados semelhantes podem ser encontrados em Baauw (2000), reportado em Schaeffer (2002), que examina o comportamento do determinante definido durante a aquisição de estruturas de posse inalienável em duas línguas distintas parametricamente, o holandês e o espanhol. Embora, na gramática alvo dessas línguas, o determinante definido exiba a possibilidade de veicular a leitura inalienável apenas em espanhol, na gramática infantil, essa vinculação é possível em ambas as línguas.

De acordo com o experimento de *Julgamento de Valor de Verdade* realizado pelo autor, envolvendo 47 crianças falantes nativas de holandês, de 4;0 a 7;0 anos de idade, e 22 adultos, também falantes nativos dessa mesma língua, observou-se que adultos e as crianças

¹⁸ Déficit Específico da Linguagem.

pertencentes ao grupo mais velho – de 6;0 a 7;0 anos – apresentam 70% de rejeição em relação à presença do determinante definido em construções inalienáveis, enquanto as crianças do grupo mais jovem – 4;0 a 5;0 anos –, embora exibam a mesma tendência, a fazem em proporções menores, apresentando 30% de rejeição em relação à mesma estrutura. Dentre os resultados do espanhol, a leitura inalienável é aceita na presença do determinante definido tanto por parte das crianças quanto por parte dos adultos.

Segundo o autor, a interpretação de determinantes definidos é afetada por dois fatores, um de ordem pragmática, em que as crianças apresentariam dificuldades para restringir o uso de determinantes aos contextos nos quais eles são permitidos na gramática alvo, e outro de ordem morfossintática, segundo o qual a aquisição incompleta dos traços morfossintáticos do núcleo D faria com que eles pudessem ser tratados como determinantes expletivos em holandês.

Pérex-Leroux *et al.* (2002a, b), que se baseiam nos estudos supracitados, realizam um estudo comparativo entre as gramáticas do AE e do espanhol, observando a interpretação de determinantes definidos no mesmo contexto possessivo. É observado que, tal qual a gramática adulta do holandês, a do AE não apresenta a possibilidade de existir a leitura inalienável com estruturas nas quais ocorre o determinante definido, como já discutido anteriormente neste capítulo, enquanto na gramática infantil pode ser observada essa veiculação.

No entanto, diferentemente de Baauw (2000), esses autores adotam a hipótese distribucional, segundo a qual “formas com distribuição morfossintática comparável competem lexicalmente por um dado espaço semântico”¹⁹ (PÉREZ-LEROUX *et al.*, 2002b, p. 199).

Assim, os autores assumem que, tanto no AE quanto no espanhol, as crianças incluiriam entre as representações semânticas do determinante definido a construção inalienável, porque haveria uma supergeneralização do uso de determinantes definidos na gramática inicial²⁰, mas não

¹⁹ Do original: “forms with comparable morphosyntactic distribution lexically compete for a given semantic space”, p. 199.

²⁰ Ver de Villiers e Roeper, 1995, Mathewson *et al.*, 2001, Pérez-Leroux *et al.*, 2004, Schaeffer e Mathewson, 2005, Munn *et al.*, 2006, entre outros, para estudos mais aprofundados sobre a aquisição de determinantes definidos em AE – comparados a outras línguas.

por dificuldades pragmáticas, já que isso seria observado mesmo em contextos discursivos e pragmáticos cuidadosamente controlados, ou por fatores morfossintáticos, dado que isso também seria notado em contextos controlados, mas pelo fato de que “eles ainda não teriam aprendido todas as restrições impostas pelas formas concorrentes em contextos particulares”²¹ (PÉREZ-LEROUX *et al.*, 2002a, p. 246).

Segundo Pérez-Leroux *et al.* (2002a), crianças falantes de línguas como o AE teriam que aprender que determinantes definidos podem ocorrer quando o referente do sintagma nominal que ele introduz pode ser identificado no discurso, é único e é máximo, tal qual em (17)²² abaixo.

(17) *The students entered the room.*

Por violarem a condição de identificação no discurso, e maximalidade, no caso de (18b), os exemplos em (18)²³, contendo nomes abstratos, como em (18a), construções de posse inalienável, como em (18b), e plurais genéricos, como em (18c), não seriam permitidos nesta língua.

- (18) a. **The freedom* depends on the development of civil society.
 b. #*The teachers* shook *the head*.
 c. #*The pandas* are not related to bears

No entanto, contextos particulares, exemplificados em (19)²⁴, como construções genéricas no singular, em (19a), e construções de posse inalienável dentro de sintagmas preposicionais, como em (19b), embora violem as mesmas condições violadas pelos exemplos acima, são permitidas em AE.

- (19) a. *The lion* lives in Africa.
 b. They were hit in *the arm*.

²¹ Do original: “they haven’t yet learned all the restrictions imposed by the competing forms in particular contexts”, p.246.

²² Pérez-Leroux *et al.*, 2002a, p. 246, exemplo 1.

²³ Pérez-Leroux *et al.*, 2002a, p. 246, exemplos 2-4.

²⁴ Pérez-Leroux *et al.*, 2002a, p. 247, exemplo 5.

Assim, crianças falantes do AE teriam que, em algum momento durante o desenvolvimento linguístico, restringir esta opção de acordo com os dados de *input*, ou seja, permitir apenas a construção alienável na presença deste determinante. Segundo Pérez-Leroux *et al.* (2002b), as evidências para isso estariam relacionadas às propriedades do sistema de possessivos do AE – que se opõem àquelas apresentadas pelo sistema do espanhol – dado que esses seriam dados mais robustos em comparação àqueles que dizem respeito às propriedades dos determinantes definidos.

Para testar a hipótese distribucional Pérez-Leroux *et al.* (2002a, b) aplicaram um experimento de compreensão (*act-out*), envolvendo, por um lado, 17 crianças falantes nativas de AE, de 3;11 a 6;6 anos de idade, e 18 adultos, como grupo controle; e por outro, 20 crianças falantes de espanhol, de 3;2 a 6;7 anos de idade, e 8 adultos, como grupo controle.

De acordo com os resultados obtidos, crianças falantes de AE pertencentes ao grupo mais jovem – nove crianças de 3;11 a 5;2 anos – permitem que a leitura inalienável seja vinculada por estruturas contendo o determinante definido mais frequentemente do que as crianças pertencentes ao grupo mais velho – oito crianças de 5;5 a 6;6 anos –, enquanto adultos não permitem essa vinculação, confirmando empiricamente a evidência assumida na pesquisa. Já em relação aos resultados do espanhol, tal qual aqueles observados em Baauw (2000), a leitura inalienável é aceita com o determinante definido tanto na gramática infantil quanto na gramática adulta.

Por fim, em relação ao PB, Mendes (2010) realiza uma pesquisa de caráter longitudinal, examinando diferentes faixas etárias de três crianças – AC, G e R – passando pela fase de aquisição dessa estrutura de posse.

De acordo com os resultados gerais dessa pesquisa, observa-se que, em 82% dos dados, o possuidor encontra-se no discurso e não na sentença, diferentemente do que propõem os estudos apresentados anteriormente – especialmente em relação àqueles que tratam de nomes de partes do corpo.

No entanto, a autora observa que, nas faixas etárias mais jovens, as crianças produzem muitos DPs soltos, o que poderia explicar essa não submissão à dependência referencial.

Por outro lado, é observado um padrão crescente com relação à presença do possuidor na sentença à medida que a criança se torna mais velha, além de, na presença do possuidor, as exigências de dependência

referencial passarem a ser respeitadas, indicando que, à medida que a gramática vai convergindo, vai sendo submetida às regras presentes na gramática adulta.

2.3 Hipóteses e previsões para a aquisição de posse inalienável

De acordo com os estudos experimentais acerca da aquisição de estruturas de posse inalienável apresentados anteriormente, a presente pesquisa assume a hipótese distribucional de Pérez-Leroux *et al.* (2002a, b, 2004), segundo a qual a superextensão do uso do determinante definido em línguas como o AE seria explicada pela restrição tardia dos contextos particulares nos quais esse determinante específico pode ocorrer nessa língua – que seriam um subconjunto dos contextos permitidos em línguas como o PB, que, por sua vez, não teriam que realizar esse mesmo tipo de restrição.

Fatores que contribuiriam para essa restrição tardia no AE constituem as propriedades das quais dependem as construções possessivas inalienáveis e seguem elencados em (I)-(II) abaixo.

- (I) Tipo da categoria vazia do possuidor: anafórica *vs* pronominal
- (II) Tipo de determinante: possessivo *vs* definido (substantivo *vs* expletivo)

Com respeito à propriedade em (I), a categoria vazia do possuidor do sintagma nominal possuído contendo o nome inalienável – conforme o seu traço [\pm lexical] (BARKER, 1995) – ou o nome que permite estabelecer uma relação inalienável – conforme o seu traço [\pm relacional] (LICHTENBERK *et al.*, 2011) – pode ser de duas naturezas, ou anafórica ou pronominal, a depender do tipo de nome envolvido/relação que ele estabelece.

Sendo um nome de parte do corpo (que estabeleça uma relação parte-todo), essa categoria é anafórica, obedecendo ao Princípio A da Teoria da Ligação, já que se trataria de um vestígio fruto de movimento, como em (20) abaixo.

- (20) a. O Pedro_j disse que o João_i lavou a mão t_i .
- b. O Pedro_j disse que os cabelos_k incomodaram o João_i.

Semelhante ao que ocorre no caso de a categoria vazia ser pronominal²⁵, tratando-se da propriedade em (II), quando o sintagma nominal que contém o nome de parte do corpo, como nos exemplos em (21), é acompanhado por um pronome possessivo, se sujeitaria ao Princípio B.

- (21) a. O Pedro_j disse que o João_i lavou a mão *dele*_{i/j/k}.
 b. O Pedro_j disse que os cabelos *dele*_{j/?i/k} incomodaram o João_i.

Em relação aos exemplos (21b), observa-se que a ligação entre o sintagma nominal possuído *os cabelos dele e João* não seria estabelecida porque não haveria c-comando.²⁶

Ainda em relação à propriedade em (II), tem-se os pronomes possessivos, que não apresentariam distinção nas leituras que podem ser vinculadas em PB ou em AE, embora apresentem diferentes estruturas sintáticas.

Além desses, tem-se os determinantes definidos, que apresentam padrões opostos em relação às leituras permitidas nessas duas línguas quando ele introduz o sintagma possuído que contém nomes potencialmente inalienáveis. Enquanto, em AE, eles seriam usados unicamente para denotar a leitura alienável, em PB, seriam usados para ambas as leituras.

De acordo com Vergnaud e Zubizarreta (1992), essa ambiguidade no PB seria explicada pelo fato de os determinantes definidos, nesta língua, se dividirem entre determinantes definidos substantivos e determinantes definidos expletivos. Enquanto os primeiros disponibilizariam a leitura alienável, os últimos disponibilizariam a leitura inalienável. E, pelo fato de o AE não apresentar determinantes expletivos – exceto por dois casos específicos já esclarecidos anteriormente e retomados mais adiante –, conseqüentemente, também não permitiria a leitura inalienável sendo vinculada por sintagmas nominais introduzidos por determinantes definidos, tal como resume o Quadro 3 abaixo.

²⁵ Quando se trata de nomes relacionais.

²⁶ Fato apontado por Grolla (c.p.).

QUADRO 3 – Leitura permitida de acordo com tipos de determinantes: AE e PB

Língua	Determinante definido		Pronome possessivo
	Substantivo	Expletivo	
PB	Leitura alienável	Leitura inalienável	Leitura (in)alienável
AE	Leitura alienável	<i>Não se aplica</i>	Leitura (in)alienável

Assumindo-se, de acordo com Vergnaud e Zubizarreta (1992), que determinantes definidos substantivos permitem a interpretação *token* do sintagma nominal e, portanto, a leitura alienável e que determinantes definidos expletivos permitem a interpretação *type* do sintagma nominal e, portanto, a leitura inalienável, assume-se que a aquisição da posse inalienável seja dependente da aquisição de determinantes substantivos e expletivos.

Assim, enquanto a gramática infantil do AE não restringe os contextos particulares em que os determinantes definidos expletivos podem aparecer – a saber, em genéricos singulares, como em (22) abaixo, chamados por Vergnaud e Zubizarreta de *protótipos*, ou seja, *tokens* que veiculam o traço [\pm espécie], e dentro de sintagmas preposicionais locativos, como em (23) abaixo –, eles seriam admitidos como substantivos e expletivos, apresentando um comportamento semelhante ao do PB.

(22) *The Bengal tiger is becoming extinct.*

(23) *John hit him on the face.*

No entanto, as evidências relevantes a favor da falta de determinantes expletivos nesta língua estariam relacionadas ao paradigma possessivo – em que os possessivos funcionam como determinantes –, como em (24) abaixo, e à impossibilidade de haver determinantes definidos introduzindo nomes próprios, como em (25).

(24) *(*The) My book is blue.*

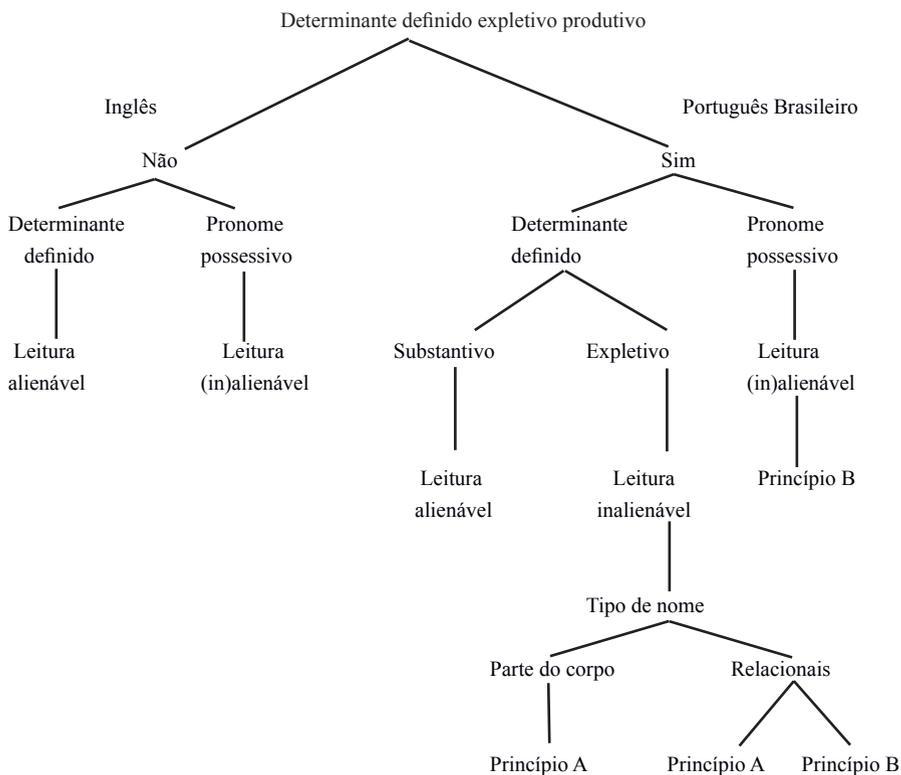
(25) *(*The) John is reading.*

Prevê-se, portanto, que crianças mais jovens falantes de AE apresentem o mesmo padrão das crianças e adultos falantes de PB, isto é, que haja uma tendência maior à aceitação de determinantes definidos coocorrendo com a leitura inalienável e uma tendência maior à aceitação de pronomes possessivos coocorrendo com a leitura alienável.

Por outro lado, crianças mais velhas falantes de AE se aproximariam do padrão apresentado por adultos falantes de AE, no qual determinantes definidos coocorrem com a leitura alienável, enquanto pronomes possessivos coocorrem com a leitura inalienável.

De acordo com o que foi discutido nesta seção, é apresentado, na Figura 1 abaixo, um esquema que resume as relações estabelecidas entre a leitura veiculada e o tipo de determinante, além de os Princípios da Ligação envolvidos, nas construções de posse em AE e PB.

FIGURA 1 – Resumo dos sistemas: PB e AE



Assim, enquanto crianças falantes de AE não restringirem os determinantes definidos expletivos aos contextos particulares em que eles são permitidos, a sua gramática estará submetida às condições da gramática de línguas como o PB.

3 Estruturas possessivas inalienáveis: viés experimental

3.1 Metodologia

Para investigar construções de posse inalienável quantificadas de forma experimental, a fim de examinar como esse tipo de estrutura se desenvolve na gramática infantil, em comparação ao seu comportamento apresentado na gramática adulta, realizou-se uma coleta transversal de dados infantis, caracterizada por Scarpa (2001, p. 204) como uma técnica de recolhimento de dados provenientes de um número alto de informantes em que as faixas etárias observadas são previamente especificadas de acordo com os achados de estudos anteriores, permitindo testar com maior profundidade certos aspectos do objeto em questão, uma vez que as estruturas estudadas dispõem de uma complexidade maior para a sua produção.

Adotando-se a técnica da *Tarefa de Julgamento de Valor de Verdade (TJVV)*, definida por Gordon (1996) como um método que “requer que a criança simplesmente faça um julgamento bipolar sobre se uma declaração descreve com precisão uma situação particular, em alusão a algum contexto ou preâmbulo” (p. 211), testa-se a compreensão de falantes de AE e de PB em relação à ocorrência de determinantes definidos e pronomes possessivos em estruturas de posse em que ambas as leituras alienável e inalienável estariam disponíveis, como será detalhado na subseção seguinte, uma vez que se espera que haja uma restrição tardia no AE dos contextos particulares em que é permitido o uso do determinante definido, fazendo com que a gramática infantil do AE se assemelhe com a gramática do PB e não com a do AE adulto.

Para a elaboração do experimento geral, foram selecionados quatro fatores, apresentados em (26) abaixo, sendo (26a-b) referentes à sentença-teste, (26c) relacionado à imagem que acompanha a sentença-teste e (26d) condizente à interação entre a sentença-teste e a imagem que a acompanha.

- (26) a. Tipo de determinante: artigo definido ou pronome possessivo
 b. Função sintática: objeto direto, sujeito da passiva, complemento do sintagma preposicional.
 c. Tipo de imagem: favorável à leitura alienável ou favorável à leitura inalienável
 d. Interpretação possessiva: alienável ou inalienável

Da interação entre os fatores (26a) e (26d), que compõem o experimento geral, resultam quatro pequenos testes, explicitados em (27) abaixo, que foram utilizados aqui para apresentar de forma didática os resultados dessa pesquisa, sendo que cada um deles contém de uma a três sentenças-teste (que coocorrem com as imagens descritas acima), e cada um dos testes interage independentemente com o fator apresentado em (26b), já que se espera ver se existe algum efeito decorrente da função sintática em que se encontra a estrutura inalienável, dado que questões de c-comando podem influenciar os resultados.

- (27) a. Definido & Inalienável
 b. Definido & Alienável
 c. Possessivo & Inalienável
 d. Possessivo & Alienável

Assim, foram conduzidos dois estudos, um em inglês e outro em português, com o mesmo material. Sendo quatro pequenos testes independentes, dois com artigos definidos (um com imagens favoráveis a leitura inalienável, outro com imagens favoráveis à leitura alienável) e dois com possessivos (também com imagens favoráveis cada tipo de leitura).

3.2 Método

Em relação aos dados infantis coletados do AE,²⁷ participaram dos testes apresentados acima 45 crianças falantes nativas, todas estudantes

²⁷ A aplicação dos testes foi permitida mediante autorização dada pela comissão de ética do Departamento de Linguística da University of Massachusetts at Amherst. Além desta, cada uma das escolas também autorizou a pesquisa e, por fim, os pais de cada uma das crianças assinaram um Termo de Consentimento Livre e Esclarecido autorizando a participação delas nos experimentos, bem como a utilização dos seus dados na presente pesquisa.

de escolas infantis²⁸ das cidades de Amherst, Sunderland e Northampton – localidades pertencentes ao estado de Massachusetts, nos EUA. Sendo o grupo de crianças mais jovens composto por 30 crianças de 3;03 a 5;11 – média etária de 4;07 – e o grupo de crianças mais velhas composto por 15 crianças de 6;01 a 8;04 – média etária de 7;02.

Além disso, como grupo controle adulto, os mesmos testes foram aplicados a 13 adultos falantes nativos de AE, sendo todos estudantes de graduação da University of Massachusetts at Amherst, EUA.

Já com respeito aos dados infantis coletados do PB²⁹, foi aplicada a mesma série de testes em 79 crianças falantes nativas, todas estudantes de escolas infantis³⁰ das cidades de Campinas, SP. Sendo o grupo de crianças mais jovens composto por 33 crianças de 3;11 a 5;10 – média etária de 4;11 – e o grupo de crianças mais velhas composto por 46 crianças de 6;00 a 9;00 – média etária de 7;06.

Além disso, como grupo controle adulto, os mesmos testes foram aplicados a 15 adultos falantes nativos de PB, sendo todos estudantes universitários da Universidade Federal de Santa Catarina e da Universidade Estadual de Campinas.

Em relação ao material, foi utilizado um computador portátil, contendo as apresentações em forma de *.ppt* pertencentes aos testes aplicados, e formulários para anotar as respostas fornecidas pelos informantes.

Assim, foram contadas três histórias em que um garoto – nas duas primeiras histórias – e uma garota – na última história – brincavam com o(a) Senhor(a) Cabeça de Batata enquanto faziam outras atividades,

²⁸ Center for Early Education and Care, em Amherst; Sunderland Elementary School, em Sunderland; e Bridge Street School, em Northampton.

²⁹ A aplicação dos testes foi permitida mediante autorização dada pelo Comitê de Ética em Pesquisa (CEP) – vide o Certificado de Apresentação para Apreciação Ética (CAAE) nº 13530014.4.0000.5404 identificando a presente pesquisa, disponível online no site da Plataforma Brasil. Além disso, cada uma das escolas também autorizou a aplicação dos testes e, por fim, os pais de cada uma das crianças assinaram um Termo de Consentimento Livre e Esclarecido autorizando a participação delas nos experimentos, bem como a utilização dos seus dados nesta pesquisa. No caso dos adultos, um Termo de Consentimento Livre e Esclarecido também foi assinado pelos próprios participantes autorizando a aplicação dos testes e a utilização dos seus dados neste trabalho.

³⁰ Escola Municipal de Educação Infantil Agostinho Páttaro, em Campinas; Escola Estadual Maria Alice Colevati Rodrigues, em Campinas.

como se lavar, se secar e comer, que também implicavam na participação do brinquedo (ou das suas partes).

Na fase dos testes, cada história contada foi ilustrada por uma série de fotografias apresentadas em *slides*, que variavam com respeito à leitura que veiculavam, bem como as sentenças-teste variavam em relação ao tipo de determinante presente.

Assim, para cada uma das sentenças-teste contendo determinante definido, como em (28), (29) e (30) abaixo, foi apresentada, por exemplo, uma fotografia veiculando a interpretação inalienável – ou seja, uma fotografia em que o personagem estivesse executando a ação numa parte do próprio corpo e não em uma parte solta do corpo do(a) Sr(a). Cabeça de Batata, que também se encontrava presente na cena –, como ilustram as Figuras 2, 3 e 4 a seguir.

- (28) a. Is Bill cleaning *the* hands?
 b. O Pedrinho está limpando *as* mãos?
- (29) a. Were *the* eyes being dried?
 b. *Os* olhos estão sendo secos?
- (30) a. Did the fly land on *the* nose?
 b. A mosca pousou *no* nARIZ?

FIGURA 2 – Objeto direto



FIGURA 3 – Sujeito da passiva



FIGURA 4 – Complemento de PP



Por outro lado, sentenças-teste contendo determinante definido também ocorreriam com uma fotografia veiculando a interpretação alienável – isto é, uma fotografia em que o personagem estivesse executando a ação numa parte do corpo solta do(a) Sr(a). Cabeça de

Batata, presente na cena, e não numa parte do próprio corpo –, como mostram os exemplos em (31), (32) e (33) abaixo, ilustrados pelas Figuras 5, 6 e 7 a seguir.

- (31) a. Is Bill cleaning *the* nose?
b. O Pedrinho está limpando *o* nariz?
- (32) a. Was *the* tongue being dried?
b. *A* língua está sendo seca?
- (33) a. Did the fly land on *the* hat?
b. A mosca pousou *no* chapéu?

FIGURA 5 – Objeto direto



FIGURA 6 – Sujeito da passiva



FIGURA 7 – Complemento de PP



Semelhante às sentenças-teste contendo determinantes definidos, as sentenças-teste contendo pronomes possessivos³¹ poderiam coocorrer com, por exemplo, uma fotografia veiculando a interpretação inalienável – na qual o personagem estivesse executando uma ação numa parte do próprio corpo e não em uma parte do corpo solta do(a) Sr(a). Cabeça de Batata –, como nos exemplos em (34), (35) e (36) abaixo, ilustrados pelas Figuras 8, 9 e 10 a seguir.

- (34) a. Is Bill cleaning *his* hand?
b. O Pedrinho está limpando a mão *dele*?

³¹ A forma *dele/dela*, no teste contendo possessivo em PB, foi utilizada no lugar de *seu/sua* porque aquela é a forma preferida na gramática adulta falada desta língua quando presentes antecedentes referenciais, como é o caso das sentenças-teste. *Seu/sua* seria a forma favorecida no caso de antecedentes genéricos.

- (35) a. Was *his* nose being dried?
b. O nariz *dele* está sendo seco?
- (36) a. Did she drop ketchup on *her* glasses?
b. Ela derramou catchup nos óculos *dela*?

FIGURA 8 – Objeto direto

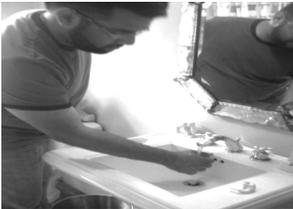


FIGURA 9 – Sujeito da passiva



FIGURA 10 – Complemento de PP



Ou ainda, as sentenças-teste contendo pronomes possessivos poderiam coocorrer com uma fotografia veiculando a interpretação alienável – na qual o personagem estivesse realizando uma ação numa parte do corpo solta do(a) Sr(a). Cabeça de Batata e não numa parte do próprio corpo –, como mostram os exemplos em (37), (38) e (39) abaixo, ilustrados pelas Figuras 11, 12 e 13 a seguir.

- (37) a. Is Bill cleaning *his* eyes?
b. O Pedrinho está limpando os olhos *dele*?
- (38) a. Was *his* ear being dried?
b. A orelha *dele* está sendo seca?
- (39) a. Did the fly land on *her* hair?
b. A mosca pousou no cabelo *dela*?

FIGURA 11 – Objeto
diretoFIGURA 12 – Sujeito da
passivaFIGURA 13 – Complemento
de PP

Aplicando essa série de experimentos, de acordo com a técnica de TJVV, foi possível verificar o julgamento do participante acerca da adequação da interpretação favorecida pela fotografia àquela veiculada pela sentença-teste, que, neste caso, se apresentava em forma de pergunta *sim-não*.

3.3 Resultados e discussão

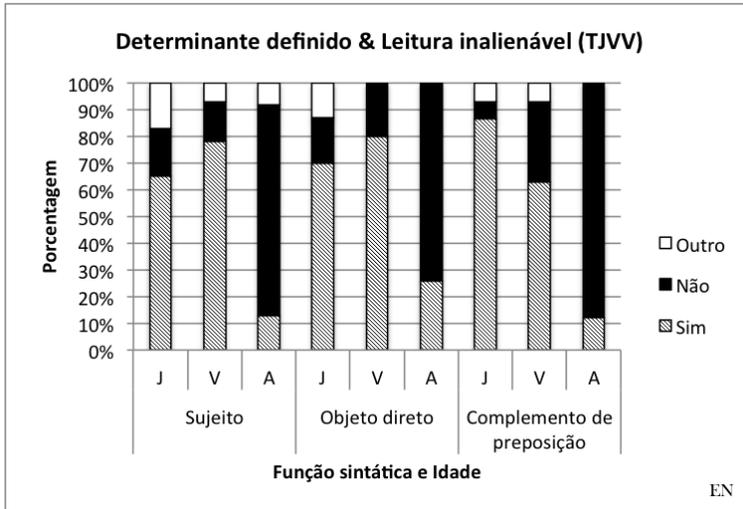
Os resultados gerais seguem de acordo com os quatro testes resultantes da combinação de fatores usados na elaboração do experimento geral, explicitadas na subseção 3.1 acima, e sua interação com o fator idade, qual seja, crianças jovens, crianças velhas e adultos, e tem por base o exame de respostas esperadas em cada um dos testes.

No *Teste 1*, o uso do determinante definido coocorrendo com este tipo de interpretação possessiva não seria permitido em AE, independentemente da função sintática sob a qual o sintagma possuído aparece.

Já em PB, essa coocorrência é permitida, desde que a categoria vazia de possuidor do sintagma possuído esteja ligada ao seu antecedente respeitando as exigências do Princípio A. Logo, espera-se que haja um resultado diferenciado em relação aos dados com o sintagma possuído desempenhando a função de sujeito da passiva.

Conforme os dados apresentados no Gráfico 1 abaixo, referente aos resultados do AE, observa-se uma taxa de respostas esperadas – no caso, *não* – bastante baixa nos resultados infantis, frente aos resultados adultos, que apresentam uma taxa mais elevada.

GRÁFICO 1 – Teste 1: AE



Legenda: J crianças mais jovens
 V crianças mais velhas
 A adultos
 EN inglês americano

Segundo o gráfico acima, a gramática das crianças mais jovens e a das mais velhas apresentam uma taxa de respostas esperadas bastante próximas – sendo de 18% e 15% na função de sujeito da passiva e 17% e 20% na função de objeto direto –, exceto quando o sintagma possuído figura na função de complemento da preposição, em que essa taxa são de 6,5% na gramática das crianças mais jovens e 30% na gramática das crianças mais velhas.

Essas diferenças percentuais são confirmadas como estatisticamente significativas conforme os resultados do Teste Exato de Fischer, exibidos na Tabela 1 abaixo.

TABELA 1 – Teste 1: valor de *p*, AE³²

Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,63240000	0,00000000
<i>Velhas</i>		-	0,00000000
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	1,00000000	0,00000001
<i>Velhas</i>		-	0,00000075
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,00771100	0,00000000
<i>Velhas</i>		-	0,00002722
<i>Adultos</i>			-

Dessa forma, quando nomes de partes do corpo coocorrem com determinantes definidos neste tipo de estrutura possessiva, observa-se que, enquanto a gramática adulta do AE permite apenas a leitura alienável, indicando o uso do determinante definido substantivo, a gramática infantil permite ambas as leituras – alienável e inalienável –, indicando o uso de ambos os tipos de determinantes definidos – substantivo e expletivo.

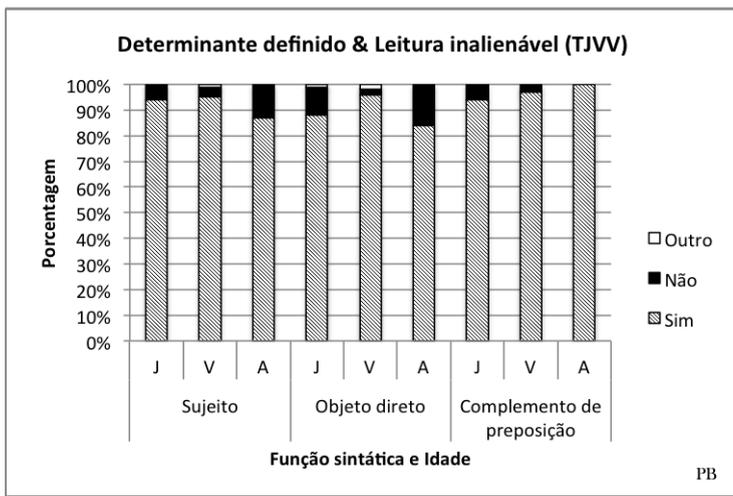
À medida que a criança fica mais velha, nota-se uma restrição com relação ao uso de um dos tipos de determinante coocorrendo com a leitura inalienável, apontando uma mudança em direção à gramática alvo, conforme mostram os resultados na função de complemento do sintagma preposicional.

Em relação aos resultados do PB adulto e infantil, observa-se, de acordo com o Gráfico 2 abaixo, uma taxa bastante alta de respostas esperadas – neste caso, *sim* – quando o sintagma possuído desempenha as funções de complemento. Porém, quando o sintagma possuído é realizado

³² As células sombreadas das tabelas contendo o valor de *p* indicam as diferenças estatisticamente significativas, de acordo com o Teste Exato de Fisher.

na função de sujeito da passiva, é notada uma baixa taxa de respostas esperadas – no caso, *não* – na gramática de todos os grupos analisados.

GRÁFICO 2 – Teste 1, PB



Legenda: J crianças mais jovens
V crianças mais velhas
A adultos
PB português brasileiro

De acordo com o gráfico acima, nas funções de complemento as taxas de respostas esperadas variam de 84% a 96% em relação aos resultados na função de objeto direto, e de 94% a 100% com respeito aos resultados na função de complemento de preposição. No entanto, quando o sintagma possuído exerce a função de sujeito da passiva, a taxa de respostas esperadas, analisando-se os três grupos de participantes, varia de 6% a 13%.

Observa-se, de acordo com a Tabela 2 abaixo, que existe uma diferença estatisticamente significativa apenas na função de sujeito da passiva e na de objeto direto – sendo esta disparidade entre a gramática das crianças mais velhas em relação a gramática adulta na primeira e na última, e entre as duas gramáticas infantis com respeito à última função sintática.

TABELA 2 – Teste 1: valor de *p*, PB

Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,53390000	0,19240000
<i>Velhas</i>		-	0,02906000
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,00229900	0,58780000
<i>Velhas</i>		-	0,00098730
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,45250000	0,30590000
<i>Velhas</i>		-	1,00000000
<i>Adultos</i>			-

Esses resultados mostram, por um lado, que a coocorrência entre esse tipo de determinante e a leitura inalienável, de fato, é permitida em todas as gramáticas examinadas, exibindo uma preferência ainda maior na gramática adulta quando o nome possuído desempenha as funções de complemento, comprovando a hipótese assumida neste estudo.

Por outro lado, os resultados concernentes à função de sujeito da passiva parecem contrariar a hipótese adotada no presente estudo, que previa não ser possível tal relação de posse, dado que, desempenhando a função de sujeito da passiva, o nome possuído não estaria ligado ao seu possuidor sob as exigências do Princípio A.

Assume-se, portanto, que, neste caso, se trata de um problema pragmático advindo da técnica experimental, em que os participantes tentaram ser colaborativos, interpretando as sentenças como “suficientemente boas”.³³

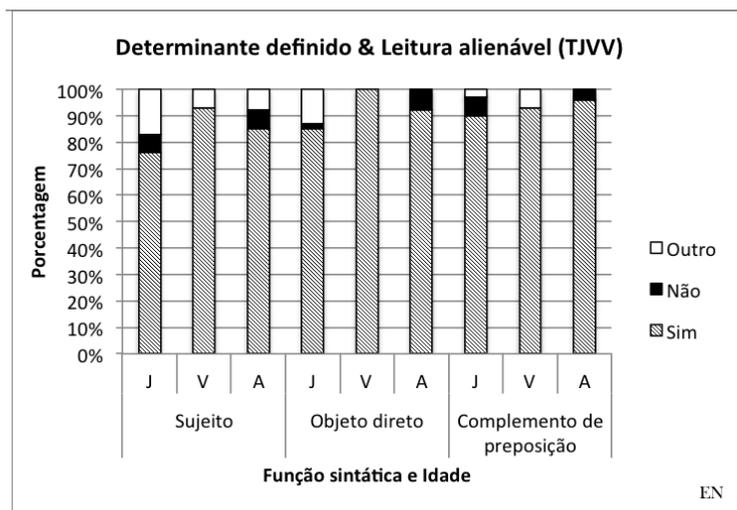
³³ *Good enough interpretation*, de acordo com Hemforth (c.p.). Neste caso, não se trata de um *good enough* das pesquisas de processamento, no sentido assumido por Collin Phillips.

Conforme o *Teste 2*, esta seria a leitura com a qual o determinante definido poderia coocorrer em AE, independentemente da função sintática sob a qual o sintagma possuído aparece.

Em PB, embora a ocorrência de determinantes seja permitida sob qualquer uma das interpretações disponíveis, sendo o determinante definido substantivo ou expletivo, haveria uma tendência a não aceitar os dados sob o *Teste 2*, quando eles desempenham as funções de complemento – em que seria preferida a leitura inalienável, tal como discutido no *Teste 1* acima. Já quando eles desempenham a função de sujeito da passiva, devido ao não cumprimento das exigências de ligação ditadas pelo Princípio A, ocorreria o oposto.

Com respeito aos resultados do AE obtidos neste teste, observa-se que, quando o determinante definido ocorre em um contexto alienável, a taxa de respostas esperadas – neste caso, *sim* – é bastante alta entre os resultados infantis e adultos, sob qualquer uma das funções sintáticas, conforme mostra o Gráfico 3 abaixo.

GRÁFICO 3 - Teste 2: AE



Segundo os resultados do gráfico acima, os dados de crianças mais jovens apresentam uma taxa de respostas esperadas que varia de 76% a 90% no presente teste, os dados de crianças mais velhas apresentam

uma variação de 93% a 100% de respostas positivas e os dados adultos variam, neste caso, de 85% a 96% de respostas *sim*.

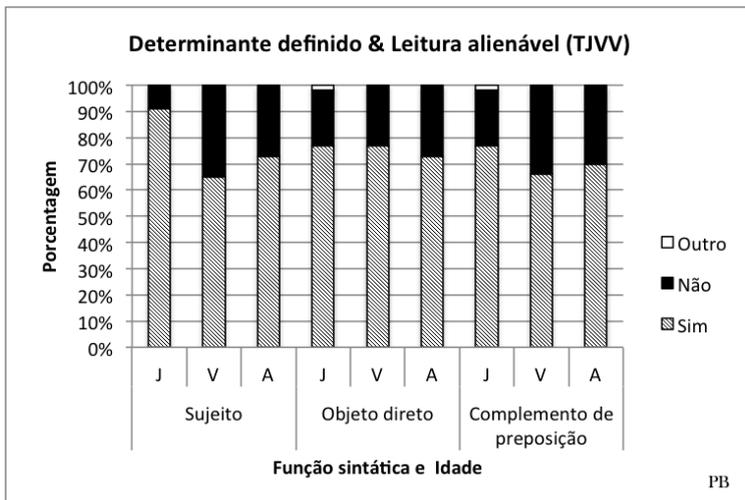
A análise estatística realizada comprova que, de fato, não há nenhuma diferença significativa entre as três gramáticas examinadas, como mostram os valores de p na Tabela 3 a seguir, comprovando a hipótese aventada nesta pesquisa.

TABELA 3 – Teste 2: valor de p , AE

Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,52770000	1,00000000
<i>Velhas</i>		-	0,46150000
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	1,00000000	0,25630000
<i>Velhas</i>		-	0,21100000
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,29890000	1,00000000
<i>Velhas</i>		-	0,48150000
<i>Adultos</i>			-

No que concerne aos resultados referentes ao PB neste teste, percebe-se que, por um lado, há uma alta taxa de respostas esperadas – no caso, *sim* – em relação aos dados na função de sujeito da passiva, e, por outro, é observada uma baixa taxa de respostas esperadas – neste caso, *não* – em relação aos dados nas funções de complemento, como mostra o Gráfico 4 abaixo.

GRÁFICO 4 – Teste 2: PB



Enquanto, na função de sujeito da passiva, a taxa de respostas esperadas varia de 65% a 91% entre os três grupos etários examinados, as taxas de respostas esperadas, nas funções de complemento, objeto direto e complemento do sintagma preposicional, variam de 21% a 27% e 21% a 30% respectivamente dentre esses mesmos três grupos.

Observa-se, de acordo com a análise quantitativa aplicada, que, de maneira geral não existe uma diferença estatisticamente significativa entre as três gramáticas examinadas, exceto em relação às gramáticas infantis quando o sintagma possuído aparece na função de sujeito da passiva, como mostra a Tabela 4 abaixo.

TABELA 4 – Teste 2: valor de *p*, PB

Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,01483000	0,18300000
<i>Velhas</i>		-	0,75360000
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	1,00000000	0,60750000
<i>Velhas</i>		-	0,80520000
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,10940000	0,44190000
<i>Velhas</i>		-	0,82420000
<i>Adultos</i>			-

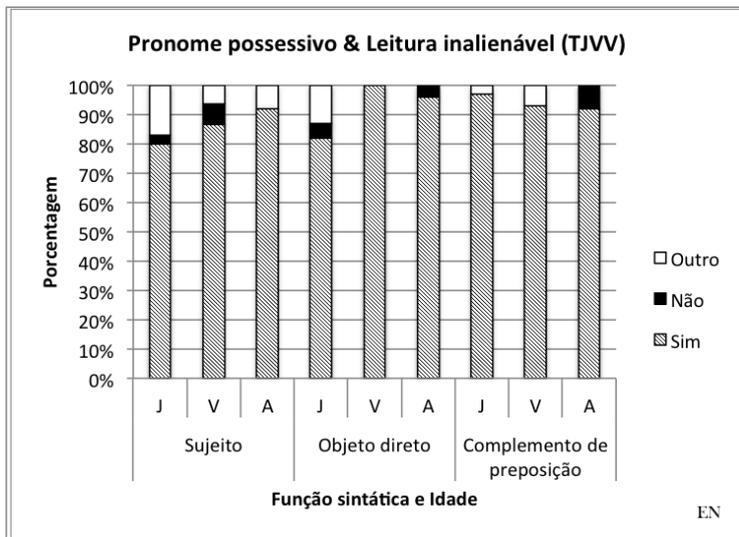
No presente teste, nota-se que crianças e adultos aceitam o uso do determinante definido sob a leitura alienável, tal como previsto para dados na função de sujeito da passiva, para o qual se apresentam altas taxas de respostas esperadas.

No entanto, não atendendo às previsões estipuladas para esta língua – já que se esperava uma preferência pela leitura inalienável mesmo em um contexto não favorecedor de tal leitura –, as gramáticas adulta e infantil também exibem uma alta taxa de aceite para esse mesmo uso quando se trata de dados na função sintática de complemento, o que não consiste em um problema, já que determinantes definidos substantivos também estão disponíveis em PB, disponibilizando tal interpretação, assim como ocorre em AE.

De acordo com o *Teste 3*, afirma-se que, embora o pronome possessivo seja permitido sob ambas as leituras, inalienável e alienável, haveria uma preferência maior pela primeira em ambas as línguas, independentemente da função sintática desempenhada, dado que, entre os possíveis possuidores, apenas um estaria saliente na sentença-teste – enquanto o outro estaria apenas presente no contexto.

Segundo os resultados do Gráfico 5 abaixo, percebe-se que, quando o pronome possessivo coocorre com a leitura inalienável em AE, a taxa de respostas esperadas – no caso, *sim* – é elevada tanto nos resultados infantis quanto nos adultos, independentemente da função sintática na qual o sintagma possuído é realizado.

GRÁFICO 5 - Teste 3: AE



Conforme o gráfico acima, a gramática das crianças mais jovens apresenta uma taxa de respostas esperadas que varia de 80% a 97%, a gramática de crianças mais velhas apresenta uma variação de 86% a 100% de respostas *sim* e a gramática adulta, de 92% a 96% de respostas positivas.

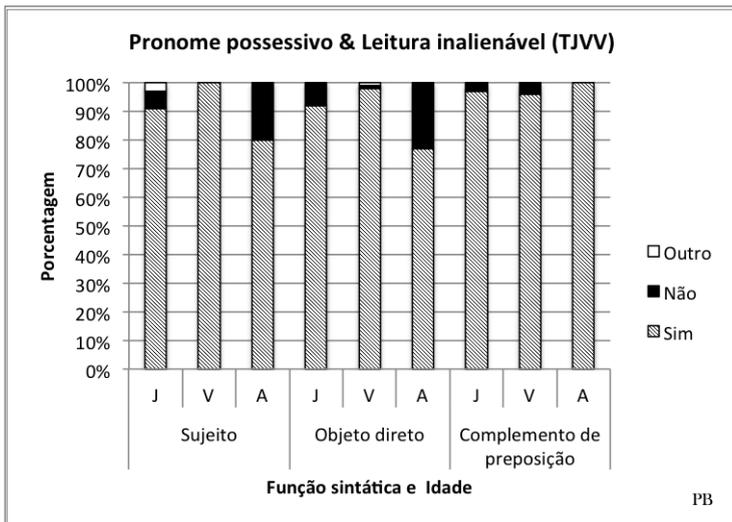
Segundo a Tabela 5 abaixo, observa-se que não há uma diferença estatística significativa entre as três gramáticas analisadas, confirmando as hipóteses da presente pesquisa.

TABELA 5 – Teste 3: valor de *p*, AE

Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	1,00000000	1,00000000
<i>Velhas</i>		-	1,00000000
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,29540000	1,00000000
<i>Velhas</i>		-	0,46430000
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	1,00000000	0,30950000
<i>Velhas</i>		-	0,48150000
<i>Adultos</i>			-

Em relação aos dados do PB, é possível observar que, embora haja uma oscilação maior entre as leituras em relação aos resultados do AE, como ilustra o Gráfico 6, a taxa de respostas esperadas – no caso, *sim* – obtidas neste teste ainda é bastante alta com respeito aos resultados infantis e adultos, independentemente da função sintática desempenhada pelo sintagma possuído.

GRÁFICO 6 – Teste 3: PB



De acordo com os resultados do gráfico acima, a taxa de respostas esperadas oriunda da gramática de crianças mais jovens varia de 91% a 97%, a de crianças mais velhas, de 96% a 100%, e a de adultos, de 77% a 100%.

Conforme a análise estatística aplicada aos dados desse teste, nota-se que existe uma discrepância significativa entre a gramática infantil e a gramática adulta nas funções de sujeito da passiva e de objeto direto, sendo a primeira entre a gramática adulta e a das crianças mais velhas e a segunda entre a adulta e ambas as gramáticas infantis, como mostra a Tabela 6 abaixo.

TABELA 6 – Teste 3: valor de *p*, PB

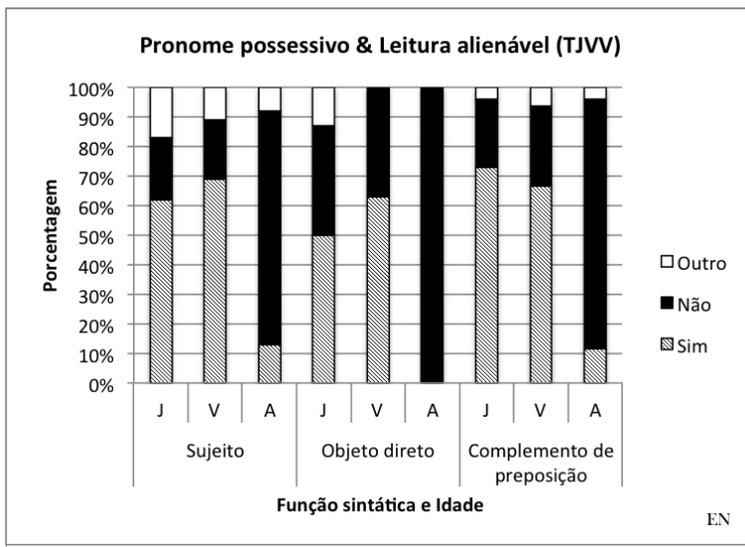
Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,16520000	0,30880000
<i>Velhas</i>		-	0,01264000
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,08318000	0,04492000
<i>Velhas</i>		-	0,00021240
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	1,00000000	1,00000000
<i>Velhas</i>		-	1,00000000
<i>Adultos</i>			-

Essa diferença estatística mostra que, entre as duas leituras disponíveis, crianças apresentam uma preferência maior pela leitura inalienável na presença do pronome possessivo do que adultos, que embora também exibam essa preferência, de acordo com os dados percentuais apresentados na tabela acima, também dispõem da leitura alienável, não favorecida pelo contexto.

Em conformidade com o *Teste 4*, tal como discutido no *Teste 3* acima, embora o uso de pronomes possessivos seja permitido sob ambas as interpretações, de acordo com a hipótese aventada neste estudo, afirma-se que haveria uma preferência maior pela leitura inalienável em ambas as línguas, independentemente da função sintática desempenhada.

No que concerne à ocorrência de pronomes possessivos sob a leitura alienável em AE, observa-se que a taxa de respostas esperadas – neste caso, *não* – segue alta na gramática adulta, enquanto se apresenta bastante inferior nos dados infantis, como mostra o Gráfico 7 abaixo.

GRÁFICO 7 – Teste 4: AE



De acordo com os resultados do gráfico acima, enquanto a gramática adulta apresenta uma taxa de respostas esperadas de 79% a 100%, a gramática infantil, de crianças mais jovens e mais velhas, apresenta uma taxa que varia de 20% a 37% de respostas negativas neste contexto.

Essa discrepância entre a gramática adulta e a gramática infantil pode ser confirmada na Tabela 7 a seguir, em que os valores de p atribuídos a cada uma delas apresenta uma diferença estatística significativa.

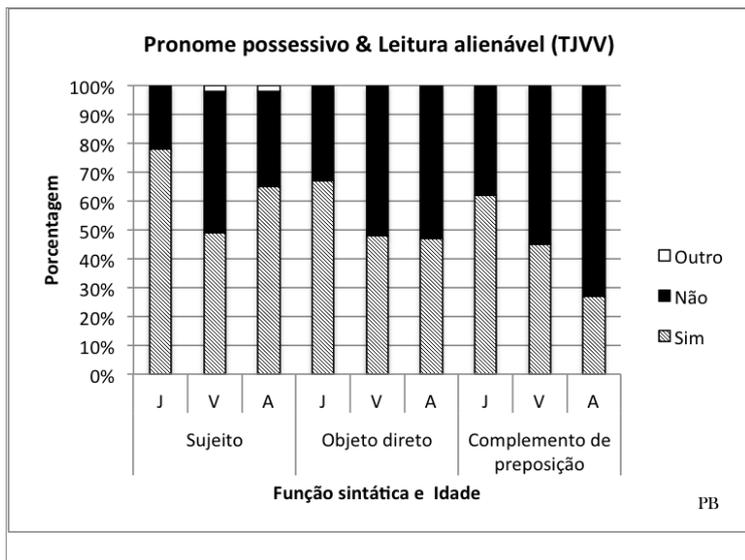
TABELA 7 – Teste 4: valor de *p*, AE

Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,82190000	0,00000000
<i>Velhas</i>		-	0,00000002
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,64780000	0,00000015
<i>Velhas</i>		-	0,00000024
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,79260000	0,00000007
<i>Velhas</i>		-	0,00001816
<i>Adultos</i>			-

Assim, em AE, quando nomes de partes do corpo coocorrem com pronomes possessivos neste tipo de construção, observa-se que, enquanto a gramática adulta dá prioridade à leitura inalienável – ainda que não seja a interpretação favorecida pelo contexto –, a gramática infantil apresenta variação entre as duas leituras possíveis, dando preferência à leitura alienável quando mais jovem e, gradualmente à leitura inalienável à medida em que vai ficando mais velha, indo em direção à gramática alvo.

Em relação aos dados do PB, é possível observar o mesmo padrão de crescimento percentual nas taxas de respostas esperadas – neste caso, *não* –, exceto em relação aos dados concernentes à função de sujeito da passiva, como mostra o Gráfico 8 abaixo.

GRÁFICO 8 – Teste 4: PB



Conforme o gráfico acima, a gramática de crianças mais jovens apresenta uma taxa de respostas esperadas variando de 22% a 38%, a gramática de crianças mais velhas exibe uma taxa que varia de 49% a 55% de respostas negativas e a gramática alvo expõe uma taxa de 33% a 73% de respostas *não*.

A análise estatística comprova que existe essa diferença significativa apenas na função de complemento da preposição – entretanto, o valor de *p* da função de objeto direto também se apresenta como bem próximo àquele esperado neste tipo de teste. No caso das outras funções, há uma diferença estatisticamente significativa apenas entre as gramáticas infantis, como mostra a Tabela 8 abaixo.

TABELA 8 – Teste 4: valor de *p*, PB

Sujeito da passiva	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,00002135	0,15100000
<i>Velhas</i>		-	0,08299000
<i>Adultos</i>			-
Objeto direto	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,02313000	0,07478000
<i>Velhas</i>		-	1,00000000
<i>Adultos</i>			-
Complemento de preposição	<i>Jovens</i>	<i>Velhas</i>	<i>Adultos</i>
<i>Jovens</i>	-	0,03610000	0,00186200
<i>Velhas</i>		-	0,09113000
<i>Adultos</i>			-

Assim, enquanto crianças mais jovens exibem uma preferência pela leitura alienável na presença de pronomes possessivos, variando de acordo com o contexto em que a sentença é apresentada, crianças mais velhas e adultos exibem uma tendência à preferência pela leitura inalienável, mesmo quando o contexto não a favorece.

4 Conclusão

Em relação aos resultados obtidos sob a técnica da TJVV, observou-se que, com respeito ao *Teste 1*, há uma baixa taxa de respostas esperadas dentre os dados infantis do AE, embora esta seja uma construção agramatical em AE adulto. Entretanto, nota-se um sensível crescimento desta ao longo das faixas etárias, demonstrando que a ocorrência do determinante definido é restrito à leitura alienável à medida que a criança fica mais velha. Os dados coletados dos sujeitos adultos exibem, por sua vez, uma taxa alta de respostas esperadas, tal como previsto.

Já em relação aos dados do PB, adultos e crianças permitem a leitura inalienável ocorrendo com o determinante definido em qualquer uma das funções sintáticas examinadas, contrariando a hipótese assumida

na pesquisa, que afirmava, neste contexto, não ser essa coocorrência preferencial quando o sintagma possuído desempenhasse a função de sujeito da passiva, já que as exigências do Princípio A seriam feridas. Assume-se que se trate, portanto, de um problema de ordem pragmática, em que os participantes seriam colaborativos com o experimento, interpretando as sentenças como “suficientemente boas”.

Quanto ao *Teste 2*, observa-se uma taxa de respostas esperadas bastante alta em AE, apontando para uma preferência pela leitura alienável neste contexto, tanto no que concerne aos dados infantis quanto no que diz respeito aos dados adultos, tal como previsto.

Em relação aos resultados do PB, percebe-se que, independentemente da função sintática exercida pelo sintagma possuído, os resultados infantis e adultos também tendem à leitura alienável, mostrando que, nesta língua, a leitura preferencial – dado que ambas estão disponíveis – varia de acordo com o contexto apresentado.

Em relação ao *Teste 3*, percebe-se que a taxa de respostas esperadas é elevada em ambas as línguas, mostrando que, neste contexto, há, de fato, uma preferência pela leitura inalienável tanto em AE quanto em PB.

Quanto ao *Teste 4*, nota-se, de maneira geral, uma alta taxa de respostas esperadas por parte dos adultos e uma baixa taxa de respostas esperadas por parte das crianças em ambas as línguas. No entanto, também é possível observar um padrão de crescimento da taxa de respostas esperadas ao longo das faixas etárias, o que mostra uma tendência à gramática-alvo.

De acordo com os resultados exibidos acima, enquanto no PB adulto haveria uma preferência pela leitura inalienável sendo veiculada por estruturas contendo determinantes definidos ou pronomes possessivos, no AE adulto, estruturas contendo pronomes possessivos favoreceriam a veiculação da leitura inalienável, enquanto estruturas contendo determinantes definidos veiculariam apenas a leitura alienável, quando presentes nomes de partes do corpo.

Ainda conforme os resultados obtidos na presente pesquisa, observa-se que no início da aquisição crianças falantes de ambas as línguas parecem aceitar ambas as interpretações sendo veiculadas por ambos os tipos de determinantes, calcando sua preferência no contexto apresentado, sendo essa preferência inicial restringida àquelas disponíveis na gramática alvo à medida que elas ficam mais velhas.

Esses resultados, portanto, corroboram as previsões aventadas na presente pesquisa pelo fato de as estruturas, por um lado poderem ser superficialmente semelhantes e, por outro, apresentarem relações mais complexas envolvendo categorias vazias distintas, além do papel sintático-semântico do determinante e sua consequente posição estrutural.

Em consequência disso, acredita-se que a hipótese assumida de que crianças falantes de AE iniciariam com uma gramática semelhante àquela apresentada pelo PB, para então, convergir para a gramática do AE, restringindo o uso de determinantes definidos, em relação a esse tipo de construção possessiva se sustenta.

Agradecimentos

À Ruth Lopes, por tanto em tanto tempo.

À CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), pelo apoio financeiro.

Referências

AUTHIER, J.-M. *The syntax of unselective binding*. 1988. Dissertation (Doctoral) – University of Southern California, Los Angeles, 1988.

BARKER, C. *Possessive descriptions*. Stanford: CSLI Publications, 1995.

BAAUW, S. *Grammatical features and the acquisition of reference*. 2000. Dissertation (Doctoral) – University of Utrecht, 2000.

BAAUW, S. Grammatical features and the acquisition of reference: a comparative study of Dutch and Spanish. Revised by Jeannette Schaeffer. *Glott International*, Blackwell Publishing, v. 6, n. 2-3, p. 65-71, Feb-Mar. 2002.

CASTRO, A. *On possessives in Portuguese*. 2006. Tese (Doutorado) – Universidade Nova de Lisboa, Université Paris 8, Vincennes Saint-Denis; Lisboa, Paris, 2006.

CASTRO, A. Sobre possessivos simples em português. In: Encontro Nacional da Associação Portuguesa de Linguística, XXII., Lisboa. *Anais...* Lisboa: APL, 2007. p. 223-237.

CHOMSKY, N. Remarks on nominalization. In: JACOBS, R.; ROSENBAUM, P. (Ed.). *Readings in English Trans-formational Grammar*. Waltham, MA: Ginn, 1970.

CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

DE VILLIERS, J.; ROEPER, T. Barriers, binding and acquisitions of NP-DP distinction. *Language Acquisition*, v. 4, n. 1-2, p. 73-104, 1995. <https://doi.org/10.1080/10489223.1995.9671660>.

FLORIPI, S. A. *Estudo da variação do determinante em sintagmas nominais possessivos na história do português*. 2008. Tese (Doutorado) – UNICAMP, Campinas, 2008.

GUÉRON, J. Inalienable possession, PRO-inclusion and lexical chains. In: GUÉRON, J.; OBENAUER, H. G.; POLLOCK, J.-Y. (Ed.). *Grammatical representation*. Dordrecht: Foris, 1985. p. 43-86.

GORDON, P. The truth-value judgment task. In: MCDANIEL, D.; MCKEE, C.; CAIRNS, H. S. *Methods for assessing children's syntax*. Cambridge, MA; London: MIT Press, 1996. p. 211-232.

LICHTENBERK, F.; VAID, J.; CHEN, H.-C. On the interpretation of alienable vs. inalienable possession: a psycholinguistic investigation. *Cognitive Linguistics*, Berlin; New York, De Gruyter, v. 22, n. 4, p. 659-689, 2011.

LONGOBARDI, G. Reference and proper names: a theory of N-movement in syntax and logical form. *Linguistic Inquiry*, MIT Press, v. 25, n. 4, p. 609-665, Fall, 1994.

MATHEWSON, L.; BRYANT, T.; ROEPER, T. A Salish stage in the acquisition of English determiners: 'unfamiliar' definites. In: SULA: THE SEMANTICS OF UNDER-REPRESENTED LANGUAGES IN THE AMERICAS, 2001, Massachusetts. *Proceedings. Occasional Papers in Linguistics*. Massachusetts: University of Massachusetts, 2001.

MENDES, F. *Estruturas possessivas: aquisição de posse funcional e posse inalienável no português brasileiro*. 2010. Dissertação (Mestrado) – Universidade Federal de Santa Catarina, Florianópolis, 2010.

MENDES, F. *Expressão da posse inalienável: um estudo diacrônico*, 2012. Manuscrito.

MUNN, A.; MILNER, K.; SCHMITT, C. Maximality and plurality in children's interpretations of definites. In: ANNUAL BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT, 30th., 2006, Somerville. *Proceedings...* Somerville: Cascadilla Press, 2006. p. 377-387.

- PÉREZ-LEROUX, A. T.; SCHMITT, C.; MUNN, A. Syntactic features and discourse factors in children's interpretation of definite determiners in inalienable possessions. In: ACL, 2002, Philadelphia. Proceedings... Philadelphia: Canadian Linguistics Association, 2002. p. 245-258.
- PÉREZ-LEROUX, A. T.; SCHMITT, C.; MUNN, A. The development of inalienable possession in English and Spanish. In: BOK-BENNEMA, R.; HOLLEBRANDSE, B.; KAMPERS-MANHE, B.; SLEEMAN, P. (Ed.). Romance languages and linguistic theory. Amsterdam: John Benjamins, 2002b. p. 199-216.
- PÉREZ-LEROUX, A. T.; SCHMITT, C.; MUNN, A.; DEIRISH, M. Learning definite determiners: genericity and definiteness in English and Spanish. In: ANNUAL BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT, 28th, 2004, Boston. *Supplement to the Proceedings...* Boston University, 2004.
- RAMOS, E. *The syntax of NPs in SLI*. 1999. Dissertation (Ph.D) – University of Massachusetts, 1999.
- RINKE, E. A combinação de artigo definido e pronome possessivo na história do português. *Estudos de Linguística Galega*, Universidade de Santiago de Compostela, v. 2, p. 121-139, 2010.
- ROBERTS, I. *Comparative syntax*. London; New York: Arnold, 1997.
- SCHAEFFER, J.; MATHEWSON, L. Grammar and pragmatics in the acquisition of article systems. *Natural Language and Linguistics Theory*, Springer, v. 23, p. 53-101, 2005. <https://doi.org/10.1007/s11049-004-5540-1>.
- TELLIER, C. *Universal licensing: implications for parasitic gap constructions*. 1988. Dissertation (Doctoral) – McGill University, Montreal, Quebec, 1988.
- VERGNAUD, J.-R.; ZUBIZARRETA, M. L. The definite determiner and the inalienable constructions in French and in English. *Linguistic Inquiry*, MIT Press, v. 23, n. 4, p. 595-652, Fall 1992.
- VIANNA, T. *O uso de artigo definido diante de pronome possessivo em textos portugueses do século XVI e XIX*. 2004. Tese (Qualificação em Sociolinguística) – UNICAMP, Campinas, 2004.

On the acquisition of distributivity in Yudja

Sobre a aquisição da distributividade em Yudja

Suzi Lima

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil

University of Toronto, Toronto, Ontario / Canada

suzilima1@gmail.com

Abstract This paper explores the acquisition of distributivity in Yudja (Tupi, Brazil) based on a study with children and adults. In this paper, we explore experimentally the hypothesis discussed in previous work (LIMA, 2008) that verbal reduplication is a pluractional marker (LASERSON, 2005) that also conveys distributivity. Two preference tasks were performed with 11 adults and 17 children. Adults' results show that the interpretation of ambiguous sentences is affected by the form of the noun phrase (conjoined nouns, pluralized nouns or bare nouns). Children's results suggest that they have an overall preference for a distributive interpretation of sentences regardless of the verb form and the NP type.

Keywords distributivity; Yudja; verbal reduplication; plural; bare nouns.

Resumo Este artigo explora a aquisição da distributividade em Yudja (Tupi, Brasil) a partir de um estudo com crianças e adultos. Nesta pesquisa, exploramos experimentalmente a hipótese discutida em trabalhos anteriores (LIMA, 2007, 2008), segundo os quais reduplicação verbal é um marcador pluracional (LASERSON, 1995) que também expressa distributividade. Duas tarefas de preferencialidade foram feitas com 11 adultos e 17 crianças. Os resultados dos adultos sugerem que a interpretação de sentenças ambíguas é afetada pelo tipo de sintagma nominal. Os resultados das crianças sugerem uma preferência por uma leitura distributiva independentemente das propriedades do verbo e do sintagma nominal.

Palavras-chave: distributividade; Yudja; reduplicação verbal; plural; nomes nus.

Recebido em: 5 de janeiro de 2017.

Aprovado em: 23 de março de 2017.

1 Introduction

Much work in formal semantics has explored the contrast between distributivity and nondistributive interpretations (collective and cumulative). One particular puzzle that will be explored in this paper is the ambiguity of sentences such as the following:

- (1) Three architects designed four buildings.
(USSERY, 1998, p. 1)

This sentence can be interpreted as distributive (covert¹ distributivity as no distributive quantifier, such as *each*, is inserted) in which case the total number of buildings is twelve. This sentence also allows nondistributive readings such as a collective reading (where the total number of buildings is four and the three architects worked together on each of them) and a cumulative reading (where the number of buildings is four, but the arrangement of the three architects working per building can vary). A debate in formal semantics is how these interpretations are derived. Authors such as Kratzer (2001) argue in favor of lexical cumulativity where collective and cumulative interpretations are derived via the same structure, and are the default interpretations. Under this view, distributivity is derived via a different operation (phrasal cumulativity): a * operator that pluralizes the VP. One prediction of this analysis is that distributivity would be more costly than nondistributive interpretations. The marginality of the distributive interpretation for sentences as (1) has been a central aspect for debate in the literature (CHAMPOLLION, to appear) and experimental evidence is inconclusive in showing whether covert distributivity is marginal for sentences such as (1) (Dotlačil, 2010 *apud* Champollion to appear). Authors such as Dotlačil (2010) argue that the marginality could be a result of a pragmatic principle: if the intention

¹ Following Champollion (to appear) we are using in this paper the terms *covert distributivity* and *overt distributivity*.

is to convey distributivity, speakers would prefer to use a distributive operator such as *each*. Below we review some studies that explore covert and overt distributivity experimentally.

Pagliarini *et al.* (2012) performed truth value judgment tasks with Italian adults and children (189 children – 4 to 13 years of age; 97 adults). In their study, for sentences like “The boys are building a tower”, adults accepted the sentence to be associated with a distributive-like picture 50% of the time. For children, an age effect was found: younger children (9 years or younger) accepted the sentence without an overt distributive quantifier as describing the distributive scene 90% of the time. Older children showed an adult-like behavior (73% acceptance for 13-year-olds). The authors argue that the rejection of the distributive interpretation of sentences with plural DPs is correlated with the acquisition of *each*.

Syrett and Mussolino (2013) have also shown, based on a truth value judgment task, that adults accept sentences with plural DPs (without *each*) to describe a distributive scene in 79% of the trials (3-year-olds did so in 91.7% of the trials and 4-year-olds in 98.3% of the trials). Note, however, that in a preference task where participants had to choose between a collective and a distributive picture (SYRETT; MUSSOLINO, 2013, Experiment 2), adults preferred a collective interpretation over a distributive one for sentences with plural DPs (without *each*) in most of the trials (88.9%). Children showed the opposite pattern (preference for collective in 31.5% of the trials). Other authors have presented similar results. Ussery (1998) has also shown a preference for non-distributive readings over distributive readings regardless the type of DP (conjoined, plural DPs) for English speaking adults. Similarly, Brooks and Braine (1996) have explored the interpretation of sentences including numerals (Three (actors) are (verb)ing an (object)) in a task where participants heard a sentence and were asked to choose the picture that was a better match (distributive or collective). Twenty adults and 100 children (4 to 9 years of age) participated in the study. The adults overwhelmingly preferred to associate a collective interpretation to sentences such as “Three children built a raft” both in the active and passive voice (97.5% and 100%, respectively). The study has also showed that the number of collective pictures selected for this kind of sentence increases with age: 4-year-olds selected the collective picture in 62.2% of the time for active voice sentences and 53.3% of the time for passive voice sentences (cf. BROOKS; BRAINE, 1996, p. 250).

Online reading time studies such as Frazier *et al.* (1999, p. 1771) have shown that distributive sentences take longer to be processed than collective sentences. According to the authors,

[...] collective interpretations may be preferred only because they avoid the need to postulate multiple events (which is claimed to be cognitively costly by Harris, Clifton, & Frazier, submitted; Majewski, in preparation, a claim termed the ‘no extra times’ or NET hypothesis). Alternatively, the simplicity of the Logical Form (LF) representation of a conjoined NP, which lacks a distribution (D) operator (Heim, Lasnik, & May, 1991), might give rise to the initial collective preference. Both hypotheses are plausible and are not mutually exclusive.

The variability of experimental results in the literature suggests that there is some cross-linguistic variation in the acceptability of covert distributivity (cf. CHAMPOLLION, to appear; 10, GIL, 1982). Despite that, most adult studies indicate that nondistributive interpretations are preferred over distributive interpretations for sentences without an overt distributive marker. Acquisition studies present a different pattern, as distributive interpretations are not necessarily dispreferred for children (cf. PAGLIARINI *et al.*, 2012; SYRETT; MUSSOLINO, 2013).

Given this debate, the goal of this paper is twofold: first, we would like to evaluate the availability of the covert distributive interpretation in Yudja; second, we want to explore the interpretation of sentences that are overtly marked for distributivity via verbal reduplication. We discuss in this paper whether the type of NP (bare singular, bare plural or conjoined NP) affects the interpretation of sentences in the case where the verb is reduplicated and in the case where it is not.

The paper is organized as follows: Section 2 presents an overview of the nominal properties of Yudja. Section 3 presents an overview of verb phrases in Yudja with a particular emphasis on the formation of verb reduplication and its meaning. Section 4 presents a formal analysis of pluractionality in Yudja. Section 5 presents an offline preference study that investigates covert and overt distributivity in the language.

2 Characteristics of Noun Phrases Yudja

Bare arguments Yudja is a number-neutral language, i.e. nouns are unspecified for number (singular, plural). Furthermore, nouns are bare and unspecified for definiteness (definite or indefinite):

- (2) *ali* *ba'i* *ixu*
 child paca eat

‘(A/the) child(ren) eat(s)/ate (a/the) paca(s)’

Literal: ‘an unspecified number of children eat an unspecified number of pacas in an unspecified number of events’.

Plural Yudja has an optional plural morpheme *-i* restricted to [+ human] nouns (FARGETTI, 2001):

- (3a) *Senahi* *kota* *ixu*
 man snake eat

‘(A/The) man/men ate/eat(s) (a/the) snake(s)’

Literal: ‘an unspecified number of men eat/ate an unspecified number of snakes in an unspecified number of events’.

- (3b) *Senahi-i* *kota* *ixu*
 man-PL snake eat

‘(The) men ate/eat (a/the) snake(s)’

Literal: ‘more than one man eat/ate an unspecified number of snakes in an unspecified number of events’.

- (3c) *kota* *senahi-i* *ixu*
 snake man-PL eat

‘(A/the) snake ate/eat(s) (the) men’

Literal: ‘an unspecified number of snakes eat/ate more than one man in an unspecified number of events’.

- (3d) * *Kota-i* *senahi* *ixu*
 snake-PL man eat

(3e) * *Senahi* *kota-i* *ixu*
 man snake-PL eat

The example (3a) means that either a single man or more than one man ate snake(s). The examples (3d) and (3e) are ungrammatical because the morpheme *-i* cannot be associated with [- human] nouns.²

3 Characteristics of verb phrases in Yudja

Verbs in Yudja are by default unspecified for number of events. The examples are presented below:

(04) *João* *ani* *ba'i* *api*
 João that paca shoot

‘João shot/shoots at that paca’

Literal: ‘João shot/shoots one single paca in an unspecified number of events’

(05) *João* *itxibi* *Maria-be* *kamema*
 João many Maria-DAT necklace

‘Pedro gave/gives many necklaces to Maria’

Literal: ‘Pedro gave/gives many necklaces to Maria in an unspecified number of events’

² Optional and restricted plural morphology for nouns is not peculiar to Yudja. Some of other Tupi languages present restrictions on the distribution of plural morphology. In Gavião (MOORE, 1984), the morpheme *éyèy* is restricted to [+animate] nouns. In Xipaya (RODRIGUES, 1995, p. 10)) the morpheme *-i* is restricted to subjects (inanimate or animate). In Kamaiurá the morphemes *-met/-het/-wet/-n* are associated with [+animate]/ [+human] (SEKY, 2000, p. 59). In Karitiana (family Arikém) and Karo (family Ramarama) there is no plural morpheme for nouns (MÜLLER; STORTO; COUTINHO, 2005; GABAS Jr., 1999, p. 41, respectively). Mekéns (family Tupari) is the only language in which an unrestricted plural morpheme is attested. The plural morpheme *-iat* is unrestricted and can be optional if other plural morphemes are overt in a sentence (GALUCIO, 2001, p. 29).

(06) *Senahi kota ixu*
 man snake eat

‘(A/the) man/men ate/eat (a/the) snake(s)’

Literal: ‘an unspecified number of men ate/eat(s) an unspecified number of snake(s) in an unspecified number of events’

In (04), (05), and (06), the number of events is strongly context dependent. In (04) and (05), a clear number of subjects (in both cases, singular subjects defined by the proper names) and a clear number of objects (in (04) a singular *paca*, in (05) a plural number of necklaces and a single beneficiary (*Maria*)) are involved. In these cases the only argument that is not specified for number is the event (*ixu* ‘eat’ and *upiku* ‘give’, respectively). Therefore, a single or many different events can be denoted by the verb, according to the context. The variety of possible readings is manifested when we have bare nouns. In (06) the two bare nouns (*senahi*, ‘man’ and *kota*, ‘snake’) are unspecified for the number of entities, i.e., they can refer to singular or plural entities. Because the verb is also unspecified for number and type of events denoted, (06) can be interpreted in a wide range of scenarios: (i) singular: one man ate one *paca*; (ii) unmarked: men ate *pacas* in various possible combinations; (iii) collective: a group of men ate (a) *paca(s)*; (iv) distributive: for each man there was an event of eating (a) *paca(s)*.

In sum, a non-reduplicated verb can have a large number of possible readings, including collective and cumulative readings. These different readings are possible because nouns in Yudja, as presented before, are number-neutral, i.e., they can denote singularities or pluralities. For that reason, the involvement of one participant or many participants in one or more than one event are possible when the noun is bare and the verb is not modified for number of events.

3.1 On the basics of reduplication: morphology

Fargetti (2001, p. 178) first described verbal reduplication in Yudja.³ The author argues that reduplication expresses both plurality of

³ Fargetti (2001) presents a phonological analysis for verbal reduplication in Yudja that I will not reproduce here because this paper focuses on the semantic aspects of verbal

arguments (subjects or objects) and reiteration (when an event is realized more than once). Fargetti argues that Yudja has two types of reduplication processes: (i) infixation (which is equivalent to partial reduplication) (*djidaku* ‘hit’ > *djidaidaku*) that expresses plurality of subjects/objects, and (ii) suffixation (which is equivalent to total reduplication) (*atxu* ‘bite’ > *atxutxu*) that expresses reiteration.

Another morphological aspect of verbal reduplication described in Lima (2008) is that inflectional morphemes do not reduplicate. This explains why in the case of the transitive verb *atxu* ‘bite’ the morpheme *a-*, which forms transitive verbs in the language, does not reduplicate. The same can be observed in a series of cases with other inflectional morphemes such as realis *-u* and verbalizers *-k-*, *-h-*. See the examples below:⁴

(07a)	Count:	a-bī T- verb root	a-bī~bī T-root~RED
(07b)	hug:	a-bīi T-verb root	a-bīi~bīi T-root~RED
(07c)	shoot	a-pī T-root	a-pī~pī T-root~RED
(07d)	extinguish	a-mi-h-u T-root-VBLZ-REALIS	a-mi~mi-h-u T-root~RED-REALIS
(07e)	hurt	du-k-u root- VBLZ-REALIS	du~du-k-u root~RED-VBLZ-REALIS
(07f)	fall	bīdī-t-u root-VBLZ-REALIS	bīdī~dī-t-u root~RED-VBLZ-REALIS

reduplication. Please consult Fargetti (2001) for a detailed analysis on the phonological properties of verbal reduplication in the language.

⁴There are exceptions for this rule such as the mood morpheme that appears reduplicated for some verbs. This aspect needs further investigation.

(07g)	peel	a-sa-k-u	a-sa~sa-k-u
		T-root-VBLZ-REALIS	T-root-RED-VBLZ-REALIS
(LIMA, 2008)			

In the examples below, Fargetti (2001) argues that partial reduplication pluralizes the subject (08b) or the object (09b) (Fargetti 2001, 178):

Partial reduplication

(08a)	<i>Una e-duku</i>	(08b)	<i>Ulu’udi</i>	<i>e-du~duku</i>
	1s REFL-hurt		1 PL	REFL-RED~hurt
	‘I hurt myself’			‘We hurt ourselves’
	(FARGETTI, 2001, p. 177)			(FARGETTI, 2001, p. 177)

Partial reduplication

(09a)	<i>Una e-djidaku</i>	<i>e-be</i>	(09b)	<i>Una ese-dji~dai~daku</i>	<i>ese-be</i>	
	1s	2s-hit	2S-DAT	1s	2PL-hit-RED	2PL-DAT
		‘I hit you’		‘I hit you (plural)’		
		(FARGETTI, 2001, p. 211)		(FARGETTI, 2001, p. 213)		

However, there are examples where partial reduplication occurs and plurality of subjects or objects cannot obtain as both refer to singular entities:

(10a)	<i>Una e-djidaku</i>	(10b)	<i>Una e-dji~dai~daku</i>
	1s 2s-hit		1s 2S-hit-RED
	‘I hit you’ (in an unspecified number of events)		‘I hit you (in different events)’

(10c)	<i>João Maria yaekua</i> João Maria remember 'João remembers/remembered Maria (unspecified number of events)'	(10d)	<i>João Maria yaekua~kua</i> João Maria remember-RED 'João always remembers/ remembered Maria'
-------	--	-------	---

Based on these facts, we argued in previous work that reduplication (partial or full reduplication) in Yudja expresses plurality of events (cf. LIMA, 2007). According to this analysis, plurality of arguments is not a result of verbal reduplication, but is due to the fact that nouns are number-neutral in Yudja. Hence, reduplication is a mechanism to pluralize the argument event and it must be dissociated from the pluralization of other arguments (subject/object).

It is not just morphological reduplication that expresses multiplicity of events or subevents. In Yudja there are verbs that are lexically reduplicated, i.e. non-reduplicated forms of these verbs are not attested. Most of the lexically reduplicated verbs describe events that have an internal repetition of events. This would explain why these verbs are lexically reduplicated. Some examples are presented below:

(11) Spoil	Lāmīmī	Fry	lusīrīrī
Sneeze	Axi'axi	Steal	pa'ia'ia
Scratch	Atxatxaku, ata'ata	blink	ipumipumi
Cry	yayaya	write	waxīwaxī
Tremble	Ari'ari	explode	pararaka
Cough	Ese'ese	blend	yukukutu
Vomit	Ena'ena	scream	azahaha
Boil	kura'kura	crawl	txitxiku

3.2 Interpretations of reduplication: semantics

In Yudja, reduplication is a way to express multiplicity of events.⁵ We will see that the distribution may be over time and/or space

⁵ In Tupi languages, verbal reduplication is a common process that is usually associated with pluractionality of events. In Mekéns (Tupari family) (GALUCIO, 2001, p. 104), any verb can be reduplicated to express iterativity or repetition of an event. In Mundurucu

or over participants. Recall examples (04) and (06) repeated below as (12a) and (13a), respectively. Unlike (12a) and (13a), their reduplicated counterparts in (12b) and (13b) cannot denote singular events:

(12a) *João* *ani* *ba'ĩ* *api*
 João that paca shoot
 ‘João shot/shoots at that paca’

Literal: ‘João shot/shoots one single paca in an unspecified number of events’

(12b) *João* *ani* *ba'ĩ* *api~pi*
 João that paca shoot.RED
 ‘João shot/shoots at that paca many times’

‘João shot/shoots at that paca in a single event’

(13a) *Senahi* *kota* *ixu*
 man snake eat
 ‘(A/the) man/men ate/eat (a/the) snake(s)’

Literal: ‘an unspecified number of men ate/eat(s) an unspecified number of snake(s) in an unspecified number of events’

(family Mundurucu) reduplication expresses the extension of the event described by the verb or its repetition (PICANÇO, 2005, p. 376; ANGOTTI, 1998, p. 15; GOMES, 2007). For Gomes (2007), reduplication may also express an extension of the duration of the event or participant pluralization. In Gavião (family Mondé) reduplication expresses repetition of an event (MOORE, 1984, p. 241). In Xipaya (family Juruna), reduplication expresses the extension of the duration of the event (RODRIGUES, 1995, p. 68). In Karitiana (family Arikém), reduplication expresses pluractionality of events (MÜLLER, SANCHEZ-MENDES, 2008). In Karo (family Ramarama) reduplication is associated with ideophones (not to verbs) and express iterativity or continuous aspect. Finally, in Kamaiurá (family Tupi Guarani) reduplication is associated with iterativity, successiveness and intensity (SEKI, 2000, p. 133-141).

- (13b) *Senahi kota ixi~ixi*
 man snake eat~RED
 ‘(A/the) man/men ate/eat (a/the) snake(s) many times’
 # ‘(A/the) man/men ate/eat (a/the) snake(s) in a single event’

According to Cusic (1981) many possible interpretations can be associated with reduplication cross-linguistically. In Yudja, one possible interpretation associated with reduplicated verbs is distributivity, as presented below (14b):

- (14a) *Wi’ubia etu*
 Tracajá egg fall
 ‘(A/the) egg(s) fell/falls down’
 Literal: ‘an unspecified number of eggs fell/falls down in an unspecified number of events’
- (14b) *Wi’ubia etu~tu*
 Tracajá egg fall-RED
 ‘(A/the) egg(s) fell/falls down (each one at a time)’

Another reading that is commonly associated with reduplication of verbs is habituality (which is a type of distributivity):

- (15a) *Una yaekua tese*
 1s remember 3PL
 ‘I remembered/remember them’
 Literal: ‘I remembered/remember them in an unspecified number of events’
- (15b) *João Maria yaekua~kua*
 João Maria remember-RED
 ‘João always remembers/remembered Maria (multiple events)’

Finally, reduplicated verbs can be associated with iterative readings:

(16a) *Akayawĩ* *i-djidaku*

door 3-hit

‘(They) hit (a/the) door’

Literal: ‘someone hit an unspecified number of doors in an unspecified number of events’

(16b) *Akayawĩ* *i-dji~dai~daku*

door 3-hit~RED

‘(They) hit (a/the) door(s) many times, continuously’

We saw that verbs can be fully or partially reduplicated in Yudja and in all cases, verbal reduplication expresses multiplicity of events. One prominent type of multiplicity of events is distributivity (in time or by participant). How multiplicity of events is interpreted ultimately depends on the context and on the Aktionsart of the verb (cf. Lima, 2007). We argue that more than one possible interpretation of the reduplicated verb is possible (either temporal distribution of participant-based distribution for instance) but in all cases the verb denotes multiple events.

3.3 Suppletion

In Yudja, suppletion is also related to plurality of events and/or arguments and the alternation between intransitive and transitive predication, similar to Xipaya (RODRIGUES, 1995), which also belongs to the Juruna family.⁶ Below, I present the verbs that have a suppletive

⁶ Like reduplication, suppletion expressing plurality is common in Tupi languages. In Mekéns (family Tupari), different verb roots denote singular/plural arguments (GALUCIO, 2001, p. 54-55). In Munduruku (family Munduruku), Dionei Moraes (p.c.) argue that there is no verbal suppletion, but only suppletion of ideophones. In Gavião (family Mondé) verbal suppletion expresses plurality of arguments or plurality of events (MOORE, 1984, p. 165, 239). In Xipaya (family Juruna), verbal suppletion has two main functions: marking valence alternation and plurality of events (RODRIGUES, 1995). The same holds of Karitiana (LUCIANA STORTO, p.c.). Finally, in Karo (family Ramarama) suppletion only occurs with intransitive roots and expresses singularity/plurality of the arguments (GABAS Jr. 1999, 46).

form. Most of the suppletive verbs in Yudja express plurality, except ‘toast’ (*u’I*, intransitive and *hunu*, transitive):

(17)		<i>Neutral verb form</i>	<i>Plural</i>
	Walk/Run	Tahu	Wãñã
	Fall	Ala	Etú
	Cry (for a death)	E’a	Yayaya
	Fly	Ãũ	Ebataku
	Paint	Itxiaku	ãpiapinu
	Get/Buy	Ita	ĩwã
	Give	(i)kua	upiku
	Drop	Daïku	daraku
	Get pregnant	Imambïu	Lapiku
	Wash	(i)tutu	aututu
	Kill	Abaku	Adikãũ
			(LIMA, 2008)

As an example, consider the verb ‘give’ which can have a neutral form as in (18a) or a plural form as in (18b):

(18a)	<i>Una</i>	<i>iidja</i>	<i>be</i>	<i>aparu</i>	<i>kua</i>
	1s	woman	DAT	beiju	give

“I gave/give beiju to (a/the) woman/women”

Literal: ‘I gave/give an unspecified number of beijus to an unspecified number of women in an unspecified number of events’.

(18b)	<i>Una</i>	<i>iidja</i>	<i>be</i>	<i>aparu</i>	<i>upiku</i>
	1s	woman	DAT	beiju	give-PL

‘I gave/give beiju(s) to each woman’

4 Theoretical analysis: reduplication and pluralization of nouns

This analysis is based on the idea that action sentences and stage-level predicates have an event argument (DAVIDSON, 1967, KRATZER, 1989). To illustrate this proposal, consider (19):

(19) The choir sang the Guillaume.

(20a) $\exists e$ [SING (e, The choir, the Marseillaise)]

(20b) $\lambda y.\lambda x.\lambda e$ SING (e, x, y)

(DAVIDSON, 1967 *apud* ROTHSTEIN, 2001)

According to Davidson (1967), the verb ‘sing’ would have three arguments: the subject (y), the object (x) and the event, the action of singing itself (e). Kratzer (1996) proposed that external arguments are not true arguments of the verb but are introduced by an additional functional head. In this view the denotation of ‘sing’ is as in (20c):

(20c) $\lambda x.\lambda e$ SING (x) (e)

Another notion that will be assumed in this paper is cumulativity. Krifka (1992) argued (see also KRATZER, 2001) that simple predicates in natural languages are typically unmarked. A predicate P is unmarked if and only if it satisfies the following formula:

(21) $\forall x\forall y: [P(x) \ \& \ P(y)] \rightarrow P(x\oplus y)$

For all x and for all y, if x is P and y is P, then the sum of x and y is P

An application of the cumulativity principle can be exemplified below, for nouns:

(22a) Cumulativity (properties of individuals) (KRATZER, 2001, p. 3 – Chapter 4, 3):

$\lambda P_{\langle et \rangle} \forall x\forall y [[P(x) \ \& \ P(y)] \rightarrow P(x+y)]]$

“A function that takes a property P of individuals and return the truth value True if for all x and for all y, if x is P and y is P, then the sum of x and y is P”

- (22b) [[child]] = {Pedro, Carlos, Maria, João, Pedro+Carlos, Pedro+Maria, Pedro+João, Carlos+Maria, Carlos+João, Maria+João, Pedro+Carlos+Maria+João}
 (based on MÜLLER; SANCHEZ-MENDES, 2008)

If we observe the extension of ‘child’ in (22b), we can see that for every two individual child (say Pedro and Maria), their sum (say Pedro+Maria) is also in the denotation of the extension of the predicate. The predicate ‘child’ is therefore unmarked. Cumulativity is also defined for verbs:

- (23a) Cumulativity (properties of events)
 (KRATZER, 2001, 3 – Chapter 4, 3):

$$\lambda P_{\langle st \rangle} \forall e \forall e' [[P(e) \& P(e')] \rightarrow P(e+e')]]$$

‘A function that takes a property P of events and return the truth value True if for all x and for all y, if x is P and y is P, then the sum of x and y is P’

- (23b) [[rain]] = {<raining₁>, {<raining₂>, ..., <raining₁+raining₂>, ...}
 (based on MÜLLER; SANCHEZ-MENDES, 2008)

If we observe the extension of ‘rain’ in (23b), we can see that for every two raining events (say raining₁ and raining₂), their sum (say raining₁+raining₂) is also in the denotation of the extension of the predicate. The predicate ‘rain’ is therefore unmarked.

As we have seen, there is no evidence that bare nouns are singular in Yudja. Rather, we have found evidence that they are unmarked, i.e. that their denotation includes both atomic individuals and sums of individuals. Consider the examples below:

- (24) João ba’i abaku
 João paca kill
 ‘João kills/killed (a/the) paca(s)’

Literal: ‘João kills/killed an unspecified number of pacas in an unspecified number events’

- (25) *Ena ba'ĩ ixu*
 2s paca eat
 'You eat/ate (a/the) paca(s)'

Literal: 'you ate/eat an unspecified number of pacas in an unspecified number of events'

- (26) *Ali ba'ĩ uatxukaha*
 child paca pursue
 '(A/the) child pursues/pursued a/the paca(s)'

Literal: 'an unspecified number of children pursued/pursues an unspecified number of pacas in an unspecified number of pacas'

In all these examples, the noun *ba'ĩ* ('paca') can be interpreted as singular or plural, definite or indefinite; its interpretation is context-dependent. This implies that these sentences are going to be true in many different scenarios where one or more pacas were killed (24), eaten (25), pursued (26) and so on. This shows that nouns in Yudja can denote a single individual or many individuals without any additional morphology.

The facts presented in Yudja support Krifka's (1992) and Kratzer's (2001) proposal, according to which predicates are unmarked from the start. Cumulativity is given at no cost for both nouns and verbs. In other words, in Yudja, lexical cumulativity can be considered the null hypothesis for any predicate (nominal or verbal) in the language (as Müller; Sanchez-Mendes (2008) proposed for Karitiana). But if verbs can denote plural events by default, what is the function of verbal reduplication?

So far we have argue that reduplicated verbs are associated with the distribution of the event over time, time-space, or participants:

Distribution over time (or time-space)

- (27) *João* *ani* *ba'i* *apĩ-pĩ*
 João that paca shoot~RED
 ‘João shot/shoots at that paca many times’
 # ‘João shot/shoots at the pace once’

Distribution over time

- (28) *Idja* *etu~tu* *da*
 woman fall~RED COL
 ‘A group of women fell/falls (each one at time)’

Participant-based distribution / Distribution over time

- (29) *Senahi* *kota* *ixi~ixi*
 man snake eat~RED
 ‘(A/the) man/men ate/eat snake(s) (more than one event)’
 # ‘(A/the) man/men ate/eat snake(s) once’

In (27), a single individual subject (João) and a single individual object (that paca) are involved in the distribution of events. As a consequence, the only possible interpretation is distribution over time. In (28), the collective morpheme *da* is inserted, and indicates that more than one group of women fell. In (29), because the nouns are number-neutral, at least two interpretations are possible: distribution of the event over time (one man ate snake(s) in different events) or participant-based distribution (more than one man ate snake(s) in different events). To explain how these readings are generated (time distribution and participant-based distribution), we will analyze reduplication as a pluractional marker (LASERSOHN, 1995).

In Lasersohn’s proposal, pluractional markers are morphemes that do not reflect the plurality of the verb’s arguments (subject/object),

but the plurality of the verb itself, because it is the verb that represents the occurrence of multiple events (LASERSOHN, 1995, p. 241). The basic denotation of pluractional markers is presented in (30a):

$$(30a) \quad V\text{-PA}(X) \Leftrightarrow \forall e \in X [P(e)] \ \& \ \text{card}(X) \geq n^7$$

'X is in the denotation of V-PA iff for all event e that belongs to X, e has the property P and the cardinality of X is at least as great as n' (LASERSOHN, 1995, p. 256)

In Lasersohn's own terms (1995, p. 240): "*pluractional markers attach to the verb to indicate a multiplicity of actions, involving multiple participants, times or locations*". That is, in Lasersohn's analysis, a verb modified by a pluractional marker can denote an event involving separate running spaces, non-overlapping running times, or even events distributed over multiple participants. In order to capture distributive readings, Lasersohn introduces a non-overlap condition in the meaning of the pluractional marker. Specifically, the condition states that for any two events e and e' , the parameters $f(e)$ and $f(e')$ of e and e' do not overlap:

$$(30b) \quad V\text{-PA}(X) \Leftrightarrow \forall e, e' \in X [P(e) \ \& \ \neg f(e) \circ f(e')] \ \& \ \text{card}(X) \geq n$$

'X is in the denotation of V-PA iff for all events e and e' in X e has a property and f(e) and f(e') do not overlap and the cardinality of X is at least as great as n' (LASERSOHN, 1995, p. 256)

Distributivity is temporal, spatio-temporal or participant-based according to the value of function f . If the distribution is temporal, f maps events to their temporal trace. If it is spatio-temporal, then f maps events to their spatio-temporal trace. Finally, if it is participant-based, f is a thematic function assigned by V , i.e. a function that maps events to a particular sort of participants, such as agents or patients.

Note that there are two possible readings of distributivity in time: the "separate" reading, which is exemplified for any example from (27) to (29) and the "continuous" reading, such as in (16b, 'Someone knocked/knocks the door many times, continuously'). In the case of "separate"

⁷ V: verb, PA: pluractional marker, e: variable over atomic events, n: variable over the natural numbers.

readings, there is a space in time between each event or subevent. To formalize this, Lasersohn included a ‘betweenness’ clause to the formula in (30b) resulting in (30c):

$$(30c) \quad V\text{-PA}(X) \Leftrightarrow \forall e, e' \in X [P(e) \& \neg f(e) \circ f(e') \& \exists x [\text{between}(x, f(e), f(e')) \& \neg \exists e'' [P(e'') \& t = f(e'')]] \& \text{card}(X) \geq n$$

‘X is in the denotation of V-PA iff for all events e, e’ in the denotation of X, e has the property P and f(e) and f(e’) do not overlap and there is an x between f(e) and f(e’) and there is no e’’ such that e’’ has the property P and t equals f(e’’) and the cardinality of x is at least as great as n’

The continuous reading, on the other hand, is generated by the negation of the betweenness clause in (30c) resulting in (30d):

$$(30d) \quad V\text{-PA}(X) \Leftrightarrow \forall e, e' \in X [P(e) \& \neg f(e) \circ f(e') \& \neg \exists x [\text{between}(x, f(e), f(e')) \& \neg \exists e'' [P(e'') \& t = f(e'')]] \& \text{card}(X) \geq n$$

‘X is in the denotation of V-PA iff for all events e, e’ in the denotation of X, e has the property P and f(e) and f(e’) do not overlap and it is not the case that there is an x between f(e) and f(e’) and there is no e’’ such that e’’ has the property P and t equals f(e’’) and the cardinality of x is at least as great as n’

Remember that Yudja is a number-neutral language. In this kind of language, a participant-based reading is possible if the noun in question is bare (31a) or plural (31b):

(31a)	<i>iidja</i>	<i>ba’i</i>	<i>ixi~ixi</i>
	woman	paca	eat-RED

‘(A/the) woman(women) eat(s)/ate (a/the) paca(s) in different events’

(31b) *iidja-i* *ba`i* *ixi~ixi*
 woman-PL paca eat~RED

‘(The) women eat(s)/ate (a/the) paca(s) in different events’

The participant-based distributive reading is not possible if nouns are singularized (by demonstratives, proper names, numerals [one]). According to Lasersohn’s proposal, if the noun is singularized, a temporal or temporal-spatial reading is forced, as we saw in (12b) (*João ani ba`i apipi* ‘João shot at that paca many times’). If the noun is plural as in (31b) or bare as in (31a), the pluractional markers allow several readings including a temporal reading (which is forced if the agent argument is singular) and a participant-based reading. In all cases, multiple events are necessarily implied.

This analysis predicts that reduplicated verbs will be associated with distributive events and that they are not just indicating plurality of events given that unmarked verbs (non-reduplicated) can also be interpreted as such. It is unclear whether the plural in nouns has the same effect (distributivity). If it does, we would expect that pluralized nouns in sentences with unmarked verbs would bias distributivity. This hypothesis is tested in the study presented in the next section.

5 Covert and overt distributivity in Yudja: a pilot study

The analysis presented in Section 3 predicts that verbal reduplication is a way to express overt distributivity. As such, we expect that when given a sentence such as (1) “Three architects designed four buildings”, if the verb is reduplicated, speakers will prefer a distributive interpretation over a nondistributive one. Conversely, if the verb is nonreduplicated, we expect that both readings will be available (distributive and nondistributive). Note, however, if the default reading is nondistributive due to lexical cumulativity, then the prediction will be that the nondistributive reading will be preferred in most trials.

One relevant question is whether the type of NP will affect the interpretation of the verb phrase. In order to test that, we manipulated three types of NPs: 1) Numeral + unmarked/bare noun (*txabiü iidja* ‘three woman’); 2) Numeral + plural (*txabiü iidjai* ‘three women’); 3) Conjoined Ns (*João Pedro Claudio djuda* ‘João, Pedro and Cláudio’). In languages

like English (USSERY, 1998), it was observed that the type of NP did not affect the results: nondistributive interpretations were preferred over distributive interpretations regardless of the NP form.⁸

Two different studies were performed in Yudja. Both studies were based on Ussery (1998)'s studies 1 and 2 for English.

Participants

A total of 17 children (5 3-to-5-year-olds; 2 7-year-olds; 5 8-to-9-year-olds; 5 10-to-11-year-olds) and 11 adults participated in this study.⁹ Participants were tested individually in a room in Yudja's local central school in the Tuba Tuba village. A local teacher familiar to the children was present to facilitate all the tasks that involved children. The study was advertised in the village amongst adults and children. Parents were individually consulted on the participation of their children in the study. The children who participated in the study volunteered to do so and were formally consented by their parents who were welcome to stay in the room along with the researcher and the research assistant.

Study 1: Conjoined Nouns (Preference task)

Materials and Methods

The study consisted of a sentence in Yudja followed by a target question. After presenting the target question, the participant was exposed to two different drawings: one corresponding to a nondistributive interpretation (collective) of the sentence and one corresponding to a distributive interpretation of the sentence. Two conditions were tested:

⁸ In previous studies with English-speaking adults and children (cf. Section 1), only pluralized DPs and conjoined DPs were manipulated as English is not a language that allows bare singulars. A similar study to the one presented in this paper in Brazilian Portuguese, a language that allows Bare Singulars, is in progress (cf. LIMA in progress).

⁹ All children who volunteered were allowed to participate (as long as their parents authorized them). Given the wide age range in the children's test group, we were not able to perform a statistical analysis that explored the age effects within this group.

Condition 1: conjoined NPs + non-reduplicated verb

(32a) *João, Pedro Cláudio djuda duwadjuse apĩ iwãnu*
 João Pedro Cláudio and four dog poison
 ‘João, Pedro and Cláudio poisoned four dogs’

‘How many dogs were poisoned?’

4

12

(nondistributive interpretation)

(distributive interpretation)

Condition 2: conjoined NPs + reduplicated verb

(32b) *João, Pedro Cláudio djuda duwadjuse apĩ iwiwãnu*
 João Pedro Cláudio and four dog poison.RED
 ‘João, Pedro and Cláudio poisoned four dogs’

‘How many dogs were poisoned?’

4

12

(nondistributive interpretation)

(distributive interpretation)

If verbal reduplication is associated with distributivity, we predict Condition 2 with reduplicated verbs to be associated with the distributive interpretation more than Condition 1 with nonreduplicated verbs.

The study included 7 critical items counterbalanced in two lists (each participant saw 4 items each).

Study 2: Bare and Plural Noun Phrases (Preference task)

Materials and Methods

The study consisted of a sentence in Yudja followed by a target question. After presenting the target question, the participant was exposed to two different drawings: one corresponding to a nondistributive interpretation (collective) of the sentence and one corresponding to

a distributive interpretation of the sentence. Four conditions were manipulated in this study:

Condition 1: bare noun + non-reduplicated verb

Condition 2: bare noun + reduplicated verb

Condition 3: plural noun + non-reduplicated verb

Condition 4: plural noun + reduplicated verb

Examples for each condition are presented below:

Condition 1: Unmarked Noun + Non-reduplicated verb

(33a) *Txabïu iidja titiri taha duwadjuse aparu ixu*
 hree woman pregnant pred four biscuit eat
 ‘Three pregnant women ate four biscuits’

Abi ne deda aparu ixu?

How many biscuit eat?

‘How many biscuits were eaten?’

4

12

(nondistributive interpretation)

(distributive interpretation)

Condition 2: Unmarked Noun + Reduplicated verb

(33b) *Txabïu iidja titiri taha duwadjuse aparu ixiixi*
 Three woman pregnant PRED four biscuit eat.RED
 ‘Three pregnant women ate four biscuits’

Abi ne deda aparu ixu?

How many biscuit eat?

‘How many biscuits were eaten?’

4

(nondistributive interpretation)

12

(distributive interpretation)

Condition 3: Plural Noun + Non-reduplicated verb

(33c) *Txabïu iidja-i tītīrī taha duwadjuse aparū iXu*

Three woman-PL pregnant PRED four biscuit eat

‘Three pregnant women ate four biscuits’

Abī ne deda aparū iXu?

How many biscuit eat?

‘How many biscuits were eaten?’

4

(nondistributive interpretation)

12

(distributive interpretation)

Condition 4: Plural Noun + Reduplicated verb

(33d) *Txabïu iidja-i tītīrī taha duwadjuse aparū ixiXi*

Three woman- PL pregnant PRED four biscuit eat.RED

‘Three pregnant women ate four biscuits’

Abī ne deda aparū iXu?

How many biscuit eat

‘How many biscuits were eaten?’

4

(nondistributive interpretation)

12

(distributive interpretation)

The study included 16 critical items counterbalanced in two lists (each participant saw 8 items each).

Results

To investigate the research questions in this study, generalized estimating equations (GEE) analyses were performed. GEE is a semi-parametric statistical technique that takes into account the interdependencies in hierarchically structured data and is appropriate for a wide variety of variable distributions (normal and skewed, continuous, dichotomous, ordinal etc.). This technique was selected to account for multiple trials for each subject within each condition of the study. Since the outcome variable in this study is binary (Collective vs distributive response), GEE analyses were performed with binomial probability distribution and logit link function.

Study 1: Results¹⁰

The percentages of Collective and Distributive responses within the two study conditions for adults and children in the sample are displayed in Table 1. As the table shows, the distribution of the two types of responses is similar across the two conditions within each age group, but the pattern is different between the age groups. Specifically, for adult subjects choose Collective more than Distributive, while children choose Distributive more than Collective.

Table 1 – Distribution of the two types of responses across conditions and age groups of the subjects

Study condition	Adult		Child	
	Collective	Distributive	Collective	Distributive
Non-Reduplicated verb	77.3%	22.7%	38.1%	61.9%
Reduplicated verb	57.1%	42.9%	42.3%	57.7%

The results of GEE analyses confirm the pattern observed in the sample. The difference between the two study conditions is nonsignificant,

¹⁰ All children that volunteered to participate in both studies were allowed to participate (as long as their parents authorized them). Given the wide age range in the children's test group (for both studies), we were not able to perform a statistical analysis that explored the age effects inside this group.

Wald $\chi^2(1) = .20$, $p = .652$. However, there is a significant difference between the two age groups, Wald $\chi^2(1) = 5.83$, $p = .016$. Based on these results, we can conclude that while the majority of adult responses (69.4% across both study conditions in the sample) are Collective – as attested in other studies presented in the literature (cf. USSERY, 1998), the majority of child responses (60.3% across both conditions in the sample) are Distributive. The pattern observed for children was also previously attested in the literature (BROOKS; BRAINE, 1996, PAGLIARINI *et al.* 2012, SYRETT; MUSSOLINO, 2013).

Study 2: Results

The percentages of Collective and Distributive responses within each of the four study conditions for groups of adults and children are displayed in Table 2. As can be seen from this table, the distribution of the two types of responses is very similar for all study conditions in the sample, with Distributive responses being more common than Collective responses. However, in the Plural + Non-Reduplicated condition the proportions of the two response types are similar for adult participants.

Table 2 – Distribution of the two types of responses across conditions and age groups of the subjects

Study condition	Adult		Child	
	Collective	Distributive	Collective	Distributive
Bare + Non-reduplicated verb	23.8%	76.2%	37.1%	62.9%
Bare + Reduplicated verb	17.4%	82.6%	36.4%	63.6%
Plural + Non-reduplicated verb	52.2%	47.8%	39.4%	60.6%
Plural + Reduplicated verb	33.3%	66.7%	37.1%	62.9%

The results of GEE analyses indicate that there is no significant difference between the age groups or study conditions in the sample and cannot be generalized beyond the sample (effect of age, Wald $\chi^2(1) = .39$, $p = .532$; effect of condition, Wald $\chi^2(3) = 4.30$, $p = .231$).

Discussion

In the first study, adults preferred a nondistributive (collective) interpretation for sentences that included conjoined nouns. Regardless of the form of the verb (reduplicated or non-reduplicated), a distributive interpretation was disfavored in most of the trials. For structures with conjoined nouns, children presented a slightly higher preference for a distributive interpretation (62%, non-reduplicated verb; 58% reduplicated verb) over a nondistributive interpretation across verb types (reduplicated or not). The difference between the age groups was statistically significant. The results found for Yudja children were consistent with the results found in previous studies in other languages (BROOKS; BRAINE, 1996; PAGLIARINI *et al.*, 2012; SYRETT; MUSSOLINO, 2013) where children favor a distributive interpretation over a collective interpretation in preference tasks, even for sentences without overt markers of distributivity (in the case of Yudja, in the absence of verbal reduplication).

In study 2, we did not find a statistically significant difference between non-reduplicated and reduplicated verbs across noun type (bare or plural) The distributive interpretation was preferred across all conditions, except for the condition with plural + Non-Reduplicated Verb.

Our initial prediction was that the distributive interpretation would be particularly favored only when the verb is reduplicated. Instead, in Study 2, we observed a preference for distributive interpretation even when the verb was not reduplicated and the noun was bare (Condition 1). We observed a preference for a distributive interpretation when the verb is reduplicated (compare Conditions 1 and 2 and Conditions 3 and 4), but this preference, as observed above, was not proven to be statistically significant.

Even though we did not find a statistically significant difference across the conditions, we observed that conditions 3 and 4 – those with pluralized nouns – had a smaller number of trials where the distributive interpretation was preferred. In the condition 3 (Bare noun + Non-reduplicated verb), in particular, there was no particular preference for collective or distributive interpretations. As such, the plural morpheme does not seem to trigger distributive interpretations.

For children, in all conditions, the results were numerically similar; based on the data presented on Table 2, there was a slight

preference for the distributive over the collective interpretation. These results are parallel to other studies with children that found a general preference for a distributive interpretation (SYRETT; MUSSOLINO, 2013).

6 Final considerations

This paper explored a formal semantics analysis of distributivity in Yudja (Tupi, Brazil). In this paper, we investigated experimentally the hypothesis discussed in previous work (LIMA, 2008) that verbal reduplication is a pluractional marker (LASERSOHN, 1995) that is also a marker of overt distributivity in the language.

Two preference tasks were performed with 11 adults and 17 children. We manipulated verb form (reduplicated and nonreduplicated) and NP type (bare noun, pluralized noun, conjoined nouns). Sentences without reduplicated verbs were taken as examples of ambiguous sentences where distributivity is covert, as in the English sentence ‘Three architects designed four buildings’ which is ambiguous between a nondistributive (collective/cumulative) and a distributive interpretation. In contrast, we hypothesized that verbal reduplication could be the manifestation of distributivity in the language.

Given that, the predictions were: 1) reduplicated verbs will trigger a preference for a distributive interpretation over a nondistributive one; 2) non-reduplicated verbs will be compatible with both interpretations (distributive and nondistributive). However, if nondistributive interpretations are favored due to lexical cumulativity (for nonreduplicated verbs), then nondistributive readings would be preferred in most trials.

The results found for Yudja children were consistent with the results found in previous studies in other languages (BROOKS; BRAINE, 1996, PAGLIARINI *et al.*, 2012; SYRETT; MUSSOLINO, 2013) where children favor a distributive interpretation over a collective interpretation in preference tasks, even for sentences without overt markers of distributivity. This was clear in Study 1 where we found a significant age effect. In Study 2, we found a numerical advantage of the distributive interpretation in all conditions, but no age effect or condition effect was found.

For adults, we found different patterns depending on the type of NP. For conjoined nouns (Study 1), we found that participants preferred the nondistributive interpretation over the distributive interpretation. The same pattern was observed in previous studies in different languages; these results are compatible with theories that claim that a non-distributive reading is the default interpretation of this kind of predicate. For pluralized nouns and bare nouns (Study 2), we found no condition effect, but the results suggest a slight preference of the distributive interpretation when the verb is reduplicated and the noun is bare (83% of the trials). In all other conditions, the preference for the distributive interpretation is lower: 76% (bare + Non-Reduplicated Verb); 67% (Plural + Reduplicated Verb) and 48% (Plural + Non-Reduplicated Verb).

By hypothesis bare nouns are more likely compatible with distributive and non-distributive interpretations because they are number-neutral and can be interpreted referring to single individuals as well as their sums. Studies on other languages that allow bare singulars in argument position are in progress and will be compared with the Yudja results (Lima in progress); also in progress is a more detailed analysis of the lexical semantics of the verbs and the interaction with overt and covert distributivity (LIMA, in progress).

Acknowledgments Funding for this research includes the Summer Research Stipend Program from UMass Amherst (2011) and Lewis and Clark Fund Grant (2009). Funai authorization (79/AEP/11).

References

- ANGOTTI, Mary Lourdes de Oliveira. *A causativização em Munduruku: aspectos morfo-sintáticos*. 1998. Dissertação (Mestrado)— University of Brasília, Brasília, 1998.
- BROOKS, Patricia J.; BRAINE, M. D. What do children know about the universal quantifiers all and each? *Cognition*, Elsevier, v. 60, n. 3, p. 235-268, 1996. [https://doi.org/10.1016/0010-0277\(96\)00712-3](https://doi.org/10.1016/0010-0277(96)00712-3)
- CHAMPOLLION, Lucas. Distributivity, collectivity, and cumulativity. Handbook article. *Wile's Companion to Semantics*. (to appear)
- CHIERCHIA, Gennaro. Mass nouns, vagueness and semantic variation. *Synthese*, Springer Link, v. 174, n. 1, p. 99-149, 2010.

- CUSIC, David. *Verbal Plurality and Aspect*. 1981. Dissertation (PhD)—Stanford University, Stanford, 1981.
- DAVIDSON, Donald. Causal Relations. *Journal of Philosophy*, Columbia University, v. 64, n. 21, p. 691-703, 1967.
- DOTLAČIL, Jakub. *Anaphora and distributivity: A study of same, different, reciprocals and others*. Utrecht, Netherlands: Utrecht University Dissertation, 2010.
- FARGETTI, Cristina. *Estudo fonológico e morfosintático da língua Juruna*. 2001. Tese (Doutorado) – Unicamp, Campinas, SP, 2001.
- FERREIRA, Marcelo. *Event Quantification and Plurality*. 1995. Dissertation (PhD) – MIT, Massachusetts, 1995.
- FRAZIER, Lyn; PACTH, Jeremy; RAYNER, Keith. Taking on semantic commitments, II: collective versus distributive readings. *Cognition*, Elsevier, v.70, p. 87-104, 1999. [https://doi.org/10.1016/S0010-0277\(99\)00002-5](https://doi.org/10.1016/S0010-0277(99)00002-5)
- GABAS Jr., Nilson. *A Grammar of Karo*. 1999. Dissertation (PhD)—University of California, Santa Bárbara, 1999.
- GALUCIO, Ana Vilacy. *The morphosyntax of Mekéns (Tupi)*. 2001. Dissertation (PhD) - University of Chicago, Chicago, 2001.
- GIL, David. Quanti er scope, linguistic variation, and natural language semantics. *Linguistics and Philosophy*, Springer Link, v. 5, n. 4, p. 421-472, 1982. <https://doi.org/10.1007/BF00355582>
- GILLON, Carrie. The mass/count distinction in Innu-aimun: implications for the meaning of plurality. In: WORKSHOP ON THE STRUCTURE AND CONSTITUENCY OF THE LANGUAGES OF THE AMERICAS (WSCLA) 15th., Vancouver, 2010. Proceedings... Vancouver: UBC Department of Linguistic, 2010.
- GOMES, Dione. Reduplicação verbal em Munduruku. In: CABRAL, Ana Suely; RODRIGUES, Aryon (Ed.). *Línguas e culturas Tupi*. Brasília: LALI/UnB, 2007. p. 391-396.
- KRATZER, Angelika. Stage-level and individual level predicates. In: *Papers in Quantification NSF Grant Report*. Amherst: Department of Linguistics, University of Massachusetts, 1989.

KRATZER, Angelika. Severing the External Argument from its Verb. In: ROORYCK, J.; ZARING, L. (Ed.). *Phrase Structure and the Lexicon*. Dordrecht: Kluwer, 1996.p. 109-137.

KRATZER, Angelika. The event argument and the semantics of verbs. Amherst: University of Massachusetts, 2001. Disponível em: <<http://semanticsarchive.net/Archive/GU1NWM4Z/>>.

KRIFKA, Manfred. Thematic relations as links between nominal reference and temporal constitution. In: SAG, Ivan A.; SAZBOLSCI, Anna (Ed.). *Lexical Matter*. Chicago: CSLI, 1992. p. 29-53.

LANDMAN, Fred. Groups, I. *Linguistics and Philosophy*, Springer Link, v. 12, n. 5, p. 559-605, 1989.

LASERSOHN, Peter. *Plurality, Conjunction, and Events*. Dordrecht, Boston: Kluwer Academic Publishers, 1995.

LIMA, Suzi. Distributivity in Brazilian Portuguese. Ms. In Progress.

LIMA, Suzi. Plurality and distributivity in Juruna: some considerations about verbal cumulativity. In: SULA, 4th., 2007, Amherst. *Proceedings...* Amherst: GLSA, 2007. p. 117-128.

LIMA, Suzi. A estrutura argumental dos verbos na língua Juruna (Yudja): da formação dos verbos para a análise das estruturas sintáticas. 2008. Dissertação (Mestrado) - Universidade de São Paulo, São Paulo, 2008.

LINK, Godehard. The logical analysis of plural and mass terms: A lattice theoretic approach. In: BÄUERLE, R.; SCHWARZE, C.; STECHOW, A. von (Ed.). *Meaning, use and interpretation of language*. Berlin: de Gruyter, 1983. p. 302-323.

MOORE, Denny. *Syntax of the language Gavião Indians of Rondônia, Brazil*. 1984. Dissertation (PhD) - University of New York, New York, 1984.

MÜLLER, Ana. The semantics of generic quantification in Brazilian Portuguese. *PROBUS*, Berlin, Mouton der Gruyter, v. 14, n. 2, 2002.

MÜLLER, Ana; STORTO, Luciana; COUTINHO, T. Number and the count-mass distinction in Karitiana. In: WORKSHOP ON THE STRUCTURE AND CONSTITUENCY OF THE LANGUAGES OF THE AMERICAS (WSCLA) 10th., Toronto, 2005. *Proceedings...* Toronto: University of Toronto, 2005.

MÜLLER, Ana; SANCHEZ-MENDES, Luciana. Pluractionality in Karitiana. In: GRØNN, Atle (Ed.). *Proceedings of SuB12*. Oslo: ILOS, 2008. p. 442-454. ISBN 978-82-92800-00-3.

PAGLIARINI, Elena; FIORIN, Gaetano; DOTLAČIL, Jakub. The acquisition of distributivity in pluralities. In: BOSTON UNIVERSITY CONFERENCE ON LANGUAGE DEVELOPMENT (BUCLD), 36th., Boston, 2012. *Proceedings...* Boston: Cascadilla Press, 2012. v. 2, p. 387-399.

PARSONS, Terence. *Events in the Semantics of English: A Study in Subatomic Semantics*, Cambridge: The MIT Press, 1990.

PICANÇO, G. L. *Munduruku: phonetics, phonology, synchrony, diachrony*. 2005. Dissertation (PhD) – University of British Columbia, 2005.

RODRIGUES, Aryon. *Línguas Brasileiras – Para o conhecimento das línguas indígenas*. São Paulo: Loyola, 1994.

RODRIGUES, Carmen L. R. *Étude morphosyntaxique de la langue Xipaya*. 1995. Dissertation (PhD) - Université Paris VII, Paris, 1995.

ROTHSTEIN, Susan. *Events and Grammar*. Netherlands: Kluwer Academic Publishers, 2001.

SAUERLAND, Uli. A new semantics for number. In: YOUNG, R.; ZHOU, Y. (Ed.). *Proceedings of SALT 13*. Ithaca, NY: CLC Publications, Cornell University, 2003.

SEKI, Lucy. *Gramática do Kamaiurá – Línguas Tupi-Guarani do alto do Xingu*. Campinas: Editora da UNICAMP, 2000.

SYRETT, Kristen; MUSOLINO, Julien. Collectivity, distributivity, and the interpretation of plural numerical expressions in child and adult language. *Language Acquisition*, Bethesda, MD, v. 20, n. 4, p. 259-291, 2013. <https://doi.org/10.1080/10489223.2013.828060>

USSERY, Cherlon. *Processing plural DPs: Collective, cumulative, and distributive interpretations*. Amherst.: University of Massachusetts, 1998.

Appendix

Glossing Abbreviations

1	first person
2	second person
3	third person
COL	collective
DAT	dative
DECL	declarative
HAB	habitual
IRR	irrealis
RED	reduplication
REFL	reflexive
S	singular
PL	plural
T	transitivizer
VBLZ	verbalizer

O processamento do numeral reduplicado *sohoji-sohoji* em Karajá: uma averiguação de ERP durante a compreensão de sentenças¹

The processing of the reduplicated numeral sohoji-sohoji in Karajá: An ERP investigation during sentence comprehension

Cristiane Oliveira da Silva

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
cristianeolivers@gmail.com

Resumo: Este trabalho tem como objetivo averiguar o processamento do numeral distributivo *sohoji-sohoji*, formado pela reduplicação do numeral *um* em Karajá, sob uma perspectiva transdisciplinar aproximando os campos da linguística teórica e da neurobiologia. O experimento de neurociência da linguagem contribui para elucidar a complexidade do processamento de sentenças com *sohoji-sohoji* quando pareado com diferentes cenários quantitativos. Utiliza-se a técnica de extração de ERP (potenciais bioelétricos relacionados a eventos) para observação e extração dos dados de latência visando a análise estatística dos dados cronométricos monitorados *on-line* por um eletroencefalograma (EEG). Os resultados das análises de variância e teste *t* demonstraram que a leitura evento-distributiva é a *default* porque a violação de exaustividade dos itens dos eventos apresentou maior complexidade de processamento do que a violação de exaustividade de participantes. Pluralidade, cardinalidade e diferenciação de eventos também apresentaram efeitos significativos durante a compreensão de sentenças com *sohoji-sohoji*.

Palavras-chave: Karajá; distributividade; ERP; processamento quantificacional.

¹ Trabalho baseado na tese de doutorado defendida na UFRJ, em 2016, sob a orientação de Marcus Maia e coorientação de Aniela Improta França.

Abstract: The aim of this paper is to investigate the processing of the reduplicated numeral *sohoji-sohoji* formed by the reduplication of the numeral *one* in Karajá, under a transdisciplinary approach, bringing together formal linguistic and neurobiology. The neuroscience of language experiment conducted sheds new light into the quantification processing complexity of *sohoji-sohoji* when different quantitative scenarios were previously presented. We use the ERPs extraction technique (Event-Related Brain Potential) to observe and to extract the wave's amplitude in order to conduct statistical analysis from the chronometric data monitored online by the electroencephalogram (EEG). The results from statistical analysis conducted (ANOVA and *t* test) suggest that the event-distributive interpretation is the first activated because the violation of event items exhaustiveness was more complex than violation of subject exhaustiveness. Plurality, cardinality and event differentiation also displayed significant effects during Karaja distributive sentence comprehension.

Keywords: Karaja; distributivity; ERP; quantification processing.

Recebido em 12 de janeiro de 2017

Aprovado em 08 de maio de 2017

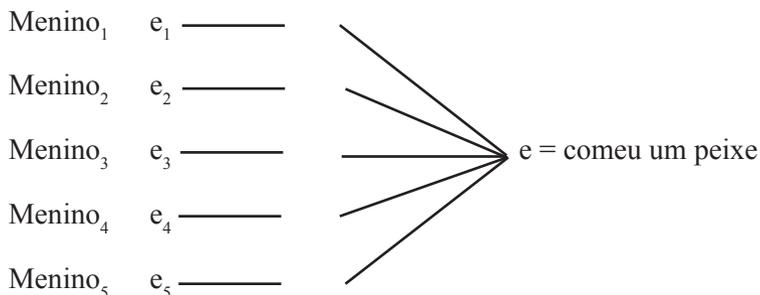
1. Introdução

A quantificação de eventos ou indivíduos é uma operação comum nas línguas naturais. Entretanto, cada sistema linguístico possui mecanismos distintos para expressar noções relacionadas à quantidade. O estudo dos quantificadores das línguas naturais é tema de interesse de diferentes áreas do conhecimento, tendo a filosofia e a lógica, por exemplo, uma gama de trabalhos clássicos que buscam averiguar o funcionamento de elementos linguísticos que exprimem quantidade. Neste trabalho, apresentaremos um estudo de processamento que utilizou a técnica de ERPs (*Event-Related Brain Potentials*) para averiguar as respostas neurobiológicas durante o processamento de sentenças quantificadas com o numeral reduplicado distributivo *sohoji-sohoji* na língua indígena Karajá.

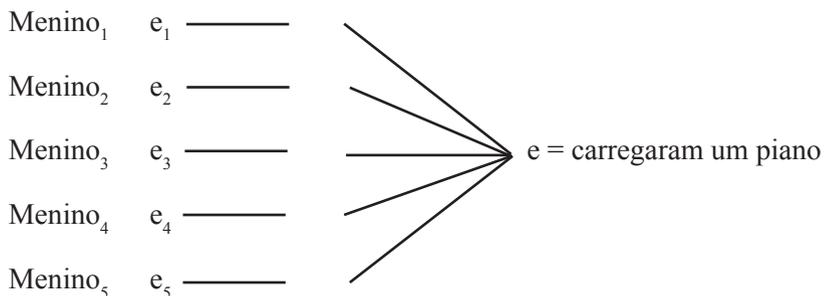
Quantificadores distributivos são responsáveis por uma operação de mapeamento entre os subeventos e os indivíduos de um enunciado.

Uma leitura distributiva prototípica deveria, portanto, operar sobre indivíduos e subeventos, sob a condição de diferenciação entre indivíduos e subeventos mapeados. Como demonstrado no exemplo (1a), para cada indivíduo x tal que x seja *menino* haverá um subevento diferente de *comer um peixe*.

(1a) Cada menino comeu um peixe.



(1b) Os meninos carregaram um piano



Contudo, delimitar determinados tipos de eventos parece não ser uma tarefa trivial e a não diferenciação dos subeventos poderia, então, resultar em uma leitura coletiva, como no exemplo (1b). Tunstall (1998), destaca ainda a existência de uma gradiência entre leituras distributivas e coletivas. Ou seja, haveria uma escala contínua entre leituras distributivas e coletivas prototípicas em que estariam disponíveis as leituras intermediárias, que seriam parcialmente distributivas ou parcialmente coletivas. Essas leituras intermediárias são tratadas por Kratzer (2003, 2007) como cumulativas.

Krifka (1998) e Kratzer (2003, 2007) argumentam a favor da cumulatividade universal dos itens lexicais simples, embora isso não garanta, de fato, que o léxico seja cumulativo na sintaxe da língua. Nessa perspectiva, assumindo que em Karajá itens lexicais sejam cumulativos também na sintaxe, a leitura cumulativa estaria sempre disponível nessa língua. Dessa forma, poderíamos especular que o processamento de uma leitura cumulativa apresentaria menor custo de processamento, exceto quando há um elemento que delimite o escopo quantificacional da sentença, como, por exemplo, um operador distributivo. Logo, ao utilizar um numeral reduplicado *sohoji-sohoji*, forçaríamos uma leitura distributiva para o enunciado, mesmo que outras leituras pudessem ser acomodadas. Neste viés, duas questões podem ser levantadas em relação à língua Karajá. (i) Seria possível processar leituras intermediárias na presença do numeral reduplicado? (ii) O pareamento de leituras intermediárias com sentenças que apresentem reduplicação numeral apresentaria maior custo de processamento? Verificamos em um estudo piloto de descrição de imagens que cenários que continham leituras intermediárias poderiam ser descritos pelos participantes utilizando o numeral *sohoji-sohoji*. Tal fato nos levou a elaborar um experimento que permitisse verificar se haveria diferenças de processamento entre as diferentes leituras quando pareadas com o numeral distributivo nessa língua.

Dessa forma, um experimento de ERPs foi aplicado para averiguar o processamento do numeral reduplicado *sohoji-sohoji* em Karajá. O experimento teve como metodologia um *design* de pareamento de imagens e sentenças (*picture matching design*). Buscamos aferir o curso do processamento das sentenças apresentadas por meio da captura dos tempos de reação e das respostas neurobiológicas dos participantes. Para tal, averiguamos o processamento de sentenças que continham o numeral reduplicado *sohoji-sohoji*, após a exposição de cenários distributivos, coletivos e parcialmente distributivos. Como veremos a seguir, cumulatividade, distributividade e cardinalidade exercem efeitos durante o processamento e interpretação de sentenças com *sohoji-sohoji* em Karajá. Tais evidências parecem dialogar com teorias vigentes acerca do processamento quantificacional e da semântica de numerais distributivos já observados em outras línguas naturais.

2. O numeral reduplicado em Karajá (e outras línguas)

A língua Karajá (família Karajá, tronco Macro-Jê) apresenta reduplicação numeral para expressar distributividade, conforme pode ser conferido no exemplo (2) abaixo. A reduplicação numeral traz à tona dois tipos de leituras distributivas, de acordo com as interpretações apresentadas em (3).

- (2) *weryry sohoji-sohoji bola-di r-e-hu-ra*
 menino um-um bola-posp 3A-VT-arremessar-PST.REC
 Lit: ‘Menino um-um arremessou bola-POSP:instrumento’

- (3) a. ‘Menino arremessou bola de um em um’ evento-distributivo
 b. ‘Cada menino arremessou (uma) bola’ participante-distributivo

Enquanto a leitura evento-distributiva (3a) implica que cada indivíduo tenha arremessado a bola em períodos de tempo distintos, a leitura participante-distributiva (3b) não apresenta essa exigência para a interpretação da sentença. Por sua vez, uma leitura coletiva parece não ser permitida, ou seja, uma leitura em que todos os meninos tenham arremessado ao mesmo tempo uma única bola.

A diferença entre as leituras acima parece ser proveniente da ambiguidade semântica engatilhada pelo numeral reduplicado *sohoji-sohoji*. Um dos objetivos deste trabalho é averiguar se durante o processamento de sentenças distributivas em Karajá haveria preferência por alguma das leituras apresentadas em (3) e se uma leitura coletiva seria imediatamente descartada, dada sua incompatibilidade com a semântica do numeral distributivo.

A possibilidade de uma leitura participante-distributiva exige que a operação de distributividade opere sobre individualidades. A impossibilidade de combinação do numeral reduplicado com nomes massivos em Karajá, que não permitem a individuação de entidades, por se tratar de substância, pode ser confirmada pela agramaticalidade do exemplo (4) abaixo. A sentença (4) demonstra que a substância *bèè* (água) é incompatível com a operação de distributividade exercida pelo operador de *sohoji-sohoji*.

- (4) *Tii bèè sohoji-sohoji-my r-e-wy-reri.
 3P água um-um-POSP 3P-VT-transportar-PRS.CONT
 Ele leva água de uma em uma/ Ele leva cada água

A reduplicação numeral como estratégia de operação distributiva parece ser relativamente comum em diversas línguas. De acordo com o site WALS,² das 251 línguas consideradas 85 apresentam a estratégia de reduplicação numeral para denotar distributividade adnominal. Gil (1982) mostra também que todas as línguas parecem dispor de recursos para expressar distributividade adverbial a partir de algum processo de morfologia numeral. Cabe ressaltar, ainda, que estes estudos abrangem pouquíssimas línguas indígenas. Inclusive, a língua Karajá, objeto de estudo deste trabalho, não faz parte do inventário de línguas observado por estes trabalhos.

Outra língua indígena que apresenta numerais reduplicados para expressar distributividade é o Karitiana (língua indígena da família Arikém). Assim como em Karajá, numerais reduplicados em Karitiana geram diferentes possibilidades de interpretação para uma mesma sentença conforme exemplificado abaixo em (5). Muller (2012) assume, a partir de testes sintáticos, que numerais distributivos em Karitiana são operadores distributivos adverbiais. Para a autora, as duas possibilidades interpretativas são o resultado da possibilidade de diferentes modos de individuação dos subeventos na denotação do predicado verbal.

- (5) Õwã nakakot sypomp.sypomp opokakosypi
 MENINO QUEBROU DOIS.DOIS OVO
 ‘Cada menino quebrou dois ovos’ – participante-distributivo
 ‘Os meninos quebraram dois ovos de cada vez’ – evento-distributivo
 Muller (2012, p. 235)

Entretanto, ao analisar a morfologia numeral em Tinglit (família Na-Dene – Canadá), Cable (2014) propõe que a distribuição numeral sempre permita que cenários participante-distributivo e evento-

² World map of languages structures online, acessado em 27 de outubro de 2014.

distributivo sejam licenciados, semelhante ao que encontramos em Karajá e Karitiana.

Entre outros fatos, eu apresentarei aqui que o local de posição sintática do numeral distributivo não tem efeito sobre o tipo de cenário distributivo que a sentença descreve. Isto é, não importa se o numeral distributivo é adverbial ou adnominal, a sentença poderá ser verdadeira tanto para cenários participante-distributivo quanto para cenários evento-distributivos.³

Cable (2014) então propõe, baseado em um teste de descrição de cenários, duas denotações semânticas, uma distributiva-adverbial, outra distributiva-adnominal para a morfologia numeral em Tinglit. Entretanto, o autor argumenta que ambas as denotações podem gerar tanto leituras distributivas sobre eventos quanto sobre indivíduos. Portanto, conclui-se que a ambiguidade gerada por numerais distributivos é recorrente nas línguas que possuem tal estratégia morfológica para expressar distributividade em seus enunciados. Com base nos dados levantados, a ambiguidade dos numerais distributivos não parece ser fruto de aspectos sintáticos e sim de características semânticas particulares a esses elementos.

Assim, com o objetivo de averiguar as possíveis diferenças de processamento entre as leituras possíveis (e impossíveis) para o numeral reduplicado em Karajá, aplicamos um experimento de EEG apresentado na seção 4 deste trabalho. Os resultados obtidos podem nos auxiliar a compreender os mecanismos cognitivos envolvidos durante o processamento da distribuição numeral. O processamento de sentenças que apresentam elementos quantificadores mostra-se um procedimento bastante complexo, uma vez que exige interface entre diferentes módulos cognitivos, como, por exemplo, o sintático e o quantificacional. É, portanto, de interesse para a neurociência cognitiva distinguir os mecanismos neurais subjacentes ao acesso a diferentes aspectos da

³ “Amongst other facts, I show here that the syntactic attachment site of the distributive numeral has no effect upon the kind of distributive scenario the sentence describes. That is, no matter whether the distributive numeral is adverbial or adnominal, the sentence may be true in either participant-distributive or event-distributive scenarios” (CABLE, 2014, p. 8)

memória semântica durante a compreensão das frases (JIANG *et al.*, 2008).

3. Pressupostos teóricos

Todas as línguas naturais conhecidas até hoje possuem mecanismos para quantificação de enunciados. Como dito na introdução, um dos objetivos deste trabalho é compreender a natureza e o processamento de um dos elementos quantificadores em Karajá, o numeral reduplicado *sohoji-sohoji*, a fim de contribuir para o entendimento de como a quantificação funciona nas línguas naturais e quais são os princípios básicos que regem estes elementos no âmbito da faculdade humana da linguagem. A capacidade de quantificar elementos linguísticos é comumente expressa por meio de operadores chamados quantificadores que, segundo a definição dada por Lyons (1977, p. 455), são modificadores que, ao se combinarem com nomes, produzem expressões cuja referência é determinada em função do tamanho do conjunto de elementos, ou em termos de quantidade da substância que está sendo referida. Entretanto, já se sabe que há também diferentes tipos de quantificadores não nominais, como é o caso, por exemplo, de certos advérbios que expressam quantidade.

A partir de uma perspectiva relacional dos quantificadores, podemos afirmar que quantificadores denotam relações binárias entre conjuntos. Os quantificadores parecem ser incumbidos da tarefa de relacionar dois conjuntos de elementos distintos, no caso, o conjunto de indivíduos e o conjunto de propriedades. Sob o viés da Semântica Formal, Heim (1982) define que um quantificador, ao relacionar dois conjuntos de elementos, forma uma tríade na qual ocupa a posição de operador da estrutura quantificada, conforme apresentado em (6) abaixo.

(6) a. Cada macaco comeu uma banana.

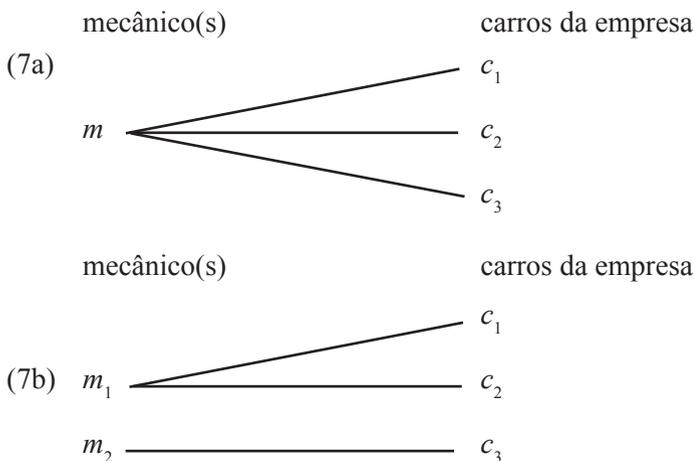
b. para	Cada _x	tal que [x é um macaco]	(\exists_y) tal que y é
	operador	sentença restritiva	escopo nuclear
	uma banana e	x comeu	y

“Para cada x tal que x é um macaco existe um y tal que y é uma banana e x comeu y”.

Portanto, do ponto de vista de análises formais, parece cabível instituir até aqui que quantificadores são operadores que executam a ação de relacionar dois conjuntos pertencentes a classes aparentemente distintas. No caso do exemplo (6), o quantificador *cada* relaciona o envolvimento dos indivíduos, que pertencem ao conjunto dos macacos, em subeventos de “comer uma banana”. Nessa área da linguística, floresceram, principalmente, análises acerca das ambiguidades de escopo causadas, sobretudo, pela interação entre sintagmas nominais quantificados.

(7) Um mecânico consertou cada carro da empresa.

A sentença (7) acima geraria pelo menos duas interpretações possíveis: (7a) há um (único) mecânico que consertou cada carro da empresa; (7b) cada carro da empresa foi consertado por um mecânico (qualquer). Essas duas interpretações podem ser esquematizadas da seguinte forma:



Sintaticamente, cada interpretação acima é representada por uma estrutura de LF (Forma Lógica) diferente que será gerada a partir do lugar de alçamento do sintagma nominal objeto. Logo, para a interpretação de sentenças com NPs quantificados em posição de objeto se faz necessário uma estrutura sintática abstrata que servirá de *input* para a interpretação semântica de sentenças com este tipo de componente estrutural (MAY, 1977, 1985).

A distributividade é uma operação que ocorre, principalmente, quando o núcleo do sintagma determinante é ocupado por um quantificador distributivo e combinado com um predicado. O predicado é, então, entendido como aplicado a cada membro individual do conjunto quantificado. Logo, a operação de um quantificador distributivo é mapear subeventos a cada átomo de indivíduos disponível no discurso. Obviamente, esta é uma definição muito simplória que não dá conta de uma série de mapeamentos distributivos que envolvem, por exemplo, ambiguidade de escopo em sentenças duplamente quantificadas. Entretanto, parece ser uma definição clara e simples o suficiente para explicar a relação entre DPs que contém quantificadores distributivos e eventos aos quais são aplicados, como em “cada macaco comeu banana”. Nesta sentença, “cada” pinça os indivíduos membros do conjunto restritor $\{x:x \text{ é macaco}\}$ sendo o predicado $[[\text{comeu banana}]]$ aplicado a cada indivíduo membro de $\{x:x \text{ é macaco}\}$. Além disso, um quantificador distributivo acaba por pluralizar o evento da sentença em que ocorrem, já que numa estrutura distributiva é exigido ao menos dois subeventos do conjunto de eventos disponíveis, ou seja, exige que a relação R entre indivíduos e eventos seja maior que 1. Logo, para a sentença (6) acima, o significado pode ser interpretado da seguinte forma: existe um evento e e para cada indivíduo m membro de $\{x:x \text{ é macaco}\}$ existe um subevento e' que é parte de e tal que em e' m “comeu banana”.

Quantificadores distributivos como *cada* do português, por exemplo, distribuem sobre entidades individuais. Portanto, ao serem aplicados a sintagmas nominais não contáveis, como aqueles ocupados por nomes massivos do tipo substância, geram sentenças agramaticais, como ilustrado em (8) abaixo, já que não seria possível individualizar suas partes, salvo sob processo de coerção ou quando uma unidade de medida é adjungida ao nome em questão. Em Karajá, o distributivo *sohoji-sohoji* exibe o mesmo padrão de comportamento, já esperado para um quantificador desta natureza, dado a gramaticalidade de (9a) e a agramaticalidade de (9b).

(8) *Ele trouxe cada água.

(9a) *Tii hãdoroó sohoji-sohoji-my r-e-wy-reri.*

3P MURICIUM-UM-PosP 3P-VT-TRANSPORTAR-PRS:CONT

‘Ele leva murici de um em um’/ ‘Ele leva cada murici’

(9b) **Tii bèè sohoji-sohoji-my rewyreri.*

3P ÁGUA UM-UM-POSP 3P-VT-TRANSPORTAR-PRS:CONT

‘Ele leva água de uma em uma’ / ‘Ele leva cada água’

Quantificadores distributivos também forçam uma leitura distributiva do evento (subeventos) da sentença em que está inserido. Portanto, quando combinados a predicados coletivos geram sentenças agramaticais (DOWTY, 1987). Por exemplo, a sentença (10) abaixo, exige que o predicado “estar acumulado no canto da sala” seja aplicado a um conjunto de indivíduos {x:x livros}, com em (10a) e (10b), e não a membros individuais deste conjunto, dada a agramaticalidade de (10c) e (10d).

(10a) Os livros estão acumulados no canto da sala.

(10b) Todos os livros estão acumulados no canto da sala.

(10c) *Um livro está acumulado no canto da sala.

(10d) *Cada livro está acumulado no canto da sala.

E, como esperado em predicados tipicamente distributivos, quantificadores distributivos podem ser aplicados normalmente, como em (11) abaixo. Da mesma forma, caso um quantificador distributivo não seja aplicado a uma sentença cujo predicado seja distributivo, ainda assim um predicado desta natureza forçaria uma leitura do tipo distributivo, mesmo não contendo um DP composto por um núcleo quantificado por um elemento distributivo em sua estrutura, como verificado em (12 e 13) abaixo.

(11) Cada menina dormiu.

(12) Todas as meninas dormiram.

(13) As meninas dormiram.

Assim, de acordo com os dados acima, parece que uma leitura distributiva deve operar não apenas sobre indivíduos > 1, mas também sobre subeventos >1, tendo, portanto, como condição a diferenciação de indivíduos e subeventos – embora seja complicado em certos predicados

delimitar subeventos. A não diferenciação de subeventos resulta em uma leitura coletiva. Entretanto, não existem apenas leituras coletivas e leituras distributivas padrão. De acordo com Tunstall (1998), as estruturas de evento distributivo-coletivas formam uma escala contínua em que leituras parcialmente-coletivas e parcialmente-distributivas estão no centro da escala e as leituras coletivas e completamente distributivas estão nas extremidades. Kratzer (2003, 2007) trata as leituras intermediárias desta escala como cumulativas. Seguindo Krifka (1998), Kratzer (2003, 2007) traz o conceito de cumulatividade como ponto crucial para a denotação de predicados em línguas naturais e argumenta a favor da propriedade de Cumulatividade Universal dos itens lexicais que pode ser conferida no trecho a seguir.

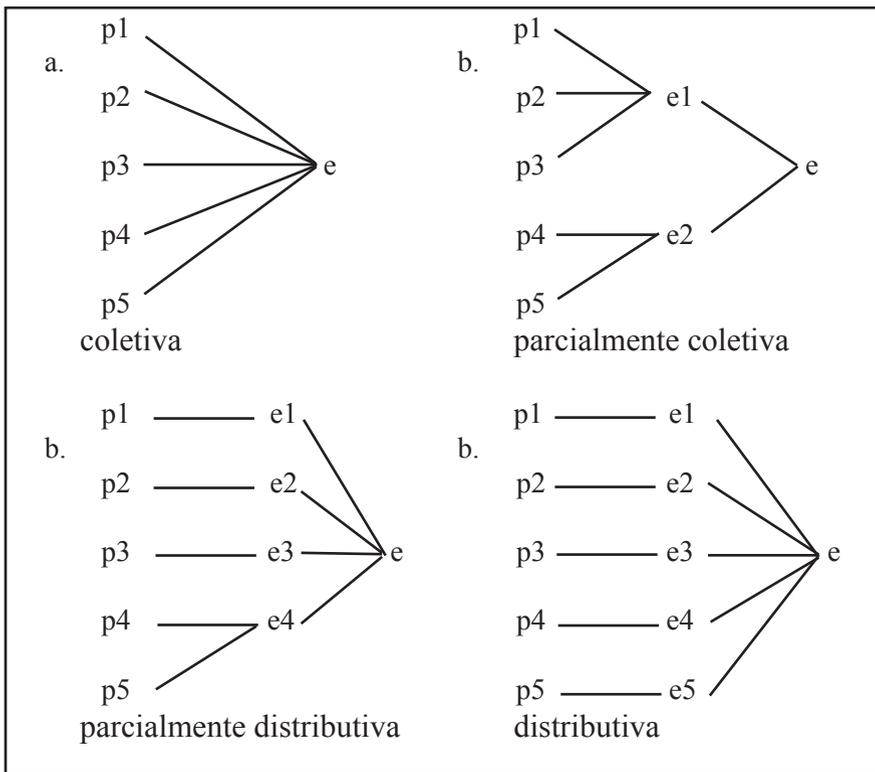
Se um predicado lógico-conceitual básico ‘vermelho’, por exemplo, é verdadeiro para meu chapéu e seu lenço (duas singularidades), então a Cumulatividade Universal diz que esse mesmo predicado também é verdadeiro para a soma de meu chapéu e seu lenço (uma pluralidade).⁴

Ou seja, em predicados simples, a leitura cumulativa estaria, portanto, disponível sem maior custo de processamento, por ser uma propriedade inerente e disponível aos itens lexicais. Entretanto, a presença de um operador distributivo resultaria em uma preferência pela leitura distributiva. Ademais, a cumulatividade inerente aos itens lexicais também explicaria o fato de que uma série de experimentos apresenta evidências de haver maior custo de processamento durante a compreensão de predicados quantificados. Isso ocorreria porque a presença de elementos quantificadores revelaria maior complexidade para o processamento semântico inicial das sentenças, que seriam cumulativas por padrão. O experimento de ERPs apresentado na próxima seção parece apontar nesta direção. Observe abaixo os diagramas abaixo, retirados de Tunstall (1998, p. 97), que buscam escrutinar as leituras intermediárias possíveis entre as leituras coletiva e distributiva padrão. Essa escala serviu para a autora demonstrar que a principal diferença entre a semântica de

⁴ “If there is a basic logical-conceptual predicate ‘red’, for example, that is true of my hat and your scarf (two singularities), then the Cumulativity Universal says that that very same predicate is also true of the sum of my hat and your scarf (a plurality).” (KRATZER, 2003, p. 12)

each e *every* diz respeito à condição de diferenciação, que é mais forte para *each* (todo subevento deverá ser distinto) do que para *every* (ao menos dois eventos deverão ser distintos).

QUADRO 1 – diagramas de leituras intermediárias de Tunstall (1998)



(14) Maria levantou todas as pastas.

Assim, para uma sentença do tipo (14) acima, vamos imaginar o seguinte cenário: há cinco pastas de documento em cima de uma mesa tal que Maria levantou todas essas pastas. Nesse cenário, há uma série de leituras possíveis de acordo com a escala acima exibida. A leitura (a) – **coletiva** – prevê que Maria levantou todas as cinco pastas de uma única vez, ou seja, houve apenas um evento de levantar as pastas. A leitura (b) – **parcialmente coletiva** – sugere que as pastas foram levantadas em pequenos grupos, por exemplo, Maria levantou três pastas e em

seguida levantou mais duas pastas gerando dois subeventos do evento de “levantar pastas”, como explícito no gráfico acima. A leitura (c) – **parcialmente distributiva** – exige que haja subeventos aplicados a membros individuais, mas permite concomitantemente subeventos coletivos, como, por exemplo, Maria levantou três pastas uma de cada vez e, depois, duas pastas ao mesmo tempo, ou seja, quatro subeventos do evento de “levantar pastas”. Por fim, a leitura (d) – **completamente distributiva** – exige que cada pasta seja levantada individualmente gerando, portanto, cinco subeventos de “levantar pasta”.

Na seção a seguir, averiguaremos, entre outras coisas, se cenários que apresentam distributividade parcial podem ser pareados com sentenças que contenham o numeral reduplicado *sohoji-sohoji* e qual seria o custo cognitivo deste procedimento. Além disso, buscamos demonstrar a aceitabilidade (ou não) de diferentes cenários quantificados quando o participante fosse exposto à uma sentença operada por *sohoji-sohoji*. Os resultados *off-line* e *on-line* corroboram para a existência de uma gradiência de aceitabilidade para diferentes leituras distributivas, na qual cenários completamente distributivos e singulares estão nas extremidades opostas da escala. Veremos adiante que, embora os resultados de testes comportamentais tenham elicitado aceitabilidade para os diferentes cenários apresentados, inclusive aqueles tipicamente não distributivos, as medidas de tempo (ms) e amplitude (μ V) parecem fornecer dados significativos para a nossa hipótese de que, embora diferentes leituras possam ser possíveis para o quantificador *sohoji-sohoji*, a leitura distributiva padrão é a primeira ativada por sentenças distributivas em Karajá cujo operador seja *sohoji-sohoji*, ou seja, a que exige menor custo cognitivo durante o processamento da sentença.

4. Processamento do numeral reduplicado *sohoji-sohoji*

Para averiguar as etapas de processamento do numeral reduplicado *sohoji-sohoji*, montamos um design experimental que permitiu a captação de dados *on-line* e *off-line*. A tarefa consistia no pareamento de diferentes cenários controlados, que eram imagens apresentadas na tela do computador, com sentenças distributivas gravadas em áudio. Este *design* é comumente conhecido como pareamento de imagens (*picture-matching design*).

O objetivo deste experimento⁵ foi observar se contextos distributivos padrão apresentariam diferenças de processamento em relação a outros contextos quantitativos (parcialmente distributivo, coletivo, não exaustivo, extra-elemento e unicidade) durante a compreensão de sentenças com o numeral reduplicado *sohoji-sohoji* em Karajá. Isto porque observamos em testagens piloto previamente averiguadas que cenários não prototípicamente distributivos eram aceitos pelos participantes como cenários possíveis para sentenças que continham *sohoji-sohoji*. Portanto, o presente experimento buscou averiguar se a aceitação de diferentes cenários quantitativos para sentenças com *sohoji-sohoji* seria resultado de um processo reflexivo envolvendo variáveis independentes como interpretação e acomodação, posteriores ao parseamento de sentenças, ou se a aceitação de diferentes cenários permitidos pelo quantificador *sohoji-sohoji* estaria relacionada às propriedades semânticas intrínsecas, geradas por este numeral reduplicado e, portanto, passíveis de serem observadas em um nível mais reflexo da compreensão de sentenças, chamado de processamento *on-line*.

4.1 Participantes

Participaram deste experimento 22 sujeitos (16 homens e 6 mulheres), com visão corrigida ou normal, audição normal, destros (OLDFIELD, 1971), falantes nativos da língua Karajá, habitantes das aldeias Btõiry ou Hãwalò (Ilha do Bananal, TO). Todos os participantes declararam não ter ingerido bebida alcoólica ou qualquer outra droga (i)lícita 24 horas antes do experimento e ter concluído o ciclo básico do ensino fundamental. A média de idade entre os participantes foi de 26 anos.

4.2 Metodologia e Materiais

O experimento consistia na apresentação da seguinte sequência de eventos: uma imagem aparecia centralizada em um monitor CRT, seguida por quatro áudios – correspondentes aos seguimentos um, dois, três e quatro de uma sentença em Karajá. Todas as sentenças experimentais

⁵ O experimento foi realizado em parceria com o laboratório ACESIN (Laboratório de Acesso Sintático coordenado por Anieli Improta França), referência na área de neurofisiologia da linguagem no Brasil.

apresentavam a estrutura sujeito-distributivo-locativo-verbo respeitando o padrão tipológico da língua, como no exemplo (15) abaixo. Após ouvir aos áudios, a tarefa dos voluntários era decidir se a imagem seria compatível, ou não, com a sentença ouvida. Este *design* é conhecido na psicolinguística experimental por *picture-matching* ou pareamento de imagens.

(15) *Krukru sohoji-sohoji òra-tyreki r-y-i-reri.*

PÁSSARO UM-UM TRONCO-POSP 3A-VT-RAIZ-PRS.CONT

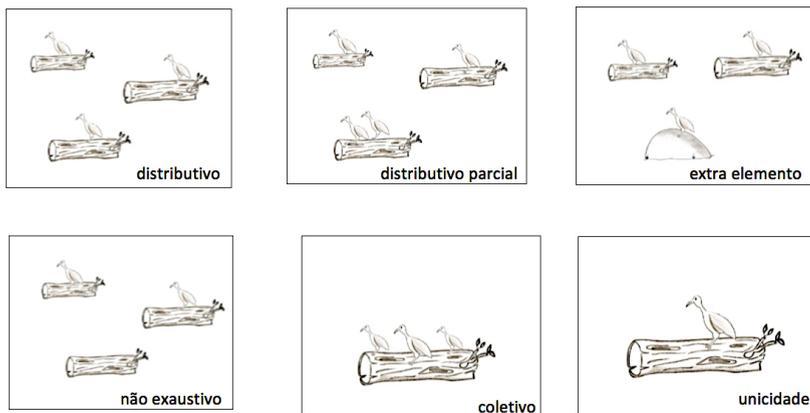
‘Cada pássaro está no tronco’

Para cada sentença experimental, foram produzidos seis cenários, a saber, (i) **distributivo padrão**, (ii) **distributivo parcial**, (iii) **extra-elemento**, (iv) **não exaustivo**, (v) **coletivo** e (vi) **unicidade**. O cenário (i) **distributivo padrão** apresentava a distributividade prototípica. Ou seja, os elementos dos conjuntos de indivíduos e as subeventualidades do conjunto de eventualidades⁶ eram distribuídos em relações de um por um, de forma que o número de relações entre os conjuntos seja maior que 2 e que haja diferenciação entre os subeventualidades. O cenário (ii) **distributivo parcial** demonstrava falha de cardinalidade na operação de relacionar indivíduos e eventualidades, uma vez que a relação de um para um era violada em uma das subeventualidades apresentadas. Os cenários (iii) **extra-elemento** e (iv) **não exaustivo** eram complementares. No primeiro, havia um elemento do grupo de indivíduos que não participava da subeventualidade descrita, já no segundo, havia uma subeventualidade que não estava relacionada a nenhum indivíduo em questão. Esses cenários testaram a propriedade de universalidade/exaustão de indivíduos e eventualidades, respectivamente. O cenário (v) **coletivo** era composto por apenas uma eventualidade para a qual todos os indivíduos eram aplicados, logo testava se propriedade de diferenciação entre subeventualidades era exigida por *sohoji-sohoji* no curso do processamento. Por fim, o cenário (vi) **unicidade** contava apenas

⁶ Em uma análise que emprega tanto predicados estativos como eventivos, o termo eventualidade (introduzido por BACH, 1986) é frequentemente utilizado como uma categorial geral que compreende ambos predicados. ‘In analyses employing both states and events, the term eventuality (introduced by BACH, 1986) is often used for the general category covering both’ (LASERSOHN, 2006).

com um átomo do grupo de indivíduos que, por sua vez, participava de uma única eventualidade, portanto, testava se um contexto singular era imediatamente descartado durante o processamento do numeral distributivo. Por exemplo, no quadro (2) abaixo podemos conferir os seis cenários gerados para a sentença (15). As imagens foram produzidas tendo como referência o banco de imagens produzidos por desenhistas Karajá no âmbito do projeto PRODOCLIN|Karajá.⁷ Cada imagem foi desenhada com lápis HB em folha A4 e digitalizada por meio de equipamento *scanner* em 300-bpi. A partir digitalizações, os cenários foram montados com o auxílio do *software* Keynote para Mac OS e exportados em formato JPG.

QUADRO 2 – Exemplos de cenários para as condições experimentais



Sentença: *Kruku sohoji-sohoji orà-tyreki ryireri*. ‘Cada pássaro está no tronco’

O experimento foi programado no software E-prime (*Psychology Software Tools, Inc. – PST*), plataforma comumente utilizada para apresentação de estímulos linguísticos com acoplamento de EEG monitorado pelo software PyCorder (*Brain Vision LLC*). Foram elaboradas seis versões do experimento em formato de quadrado latino, de

⁷ Projeto de Documentação da Língua e Cultura Karajá, parte do Programa de Documentação de Línguas Indígenas do Museu do Índio (FUNAI), apoiado pela UNESCO.

maneira que nenhum participante viu as mesmas sentenças experimentais, porém todos os participantes viram todas as condições experimentais a partir de diferentes sentenças.

As listas foram apresentadas tanto na ordem normal como inversa, para evitar efeitos de ordem de lista em cada versão. A partir da criação das versões experimentais distribuídas entre sujeitos – *design* de quadrado latino – é possível garantir que um mesmo sujeito nunca veja as mesmas sentenças, mas seja exposto a sentenças de todas as condições. Ao todo, para este experimento, seis listas/versões foram programadas de acordo com a distribuição dos materiais experimentais em formato de quadrado latino. Este protocolo experimental é importante para inibir que fatores extralinguísticos como memória, por exemplo, atuem de modo a interferir no processo de compreensão de sentenças. A distribuição dos materiais entre os participantes também auxilia para que os dados obtidos sejam mais robustos, uma vez que a análise entre sujeitos dilui fatores subjetivos relativos aos indivíduos. Além disso, os estímulos foram pseudo-randomizados dentro de cada lista de cada versão, seguindo as seguintes restrições: (i) estímulos experimentais foram separados por, pelo menos, um item distrator; (ii) nenhuma sentença experimental foi seguida por outra da mesma condição.

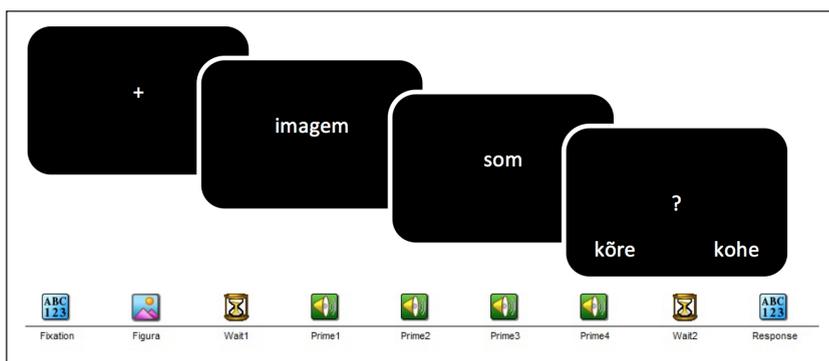
Formam o conjunto total de materiais deste experimento 48 sentenças experimentais acrescidas de 48 distratoras gravadas em áudio e 336 imagens digitalizadas e tratadas. Os áudios foram traduzidos do português para o Karajá em conjunto com dois professores de língua Karajá, conferidos por um terceiro falante, e, finalmente, gravados por um consultora Karajá. Cada sentença experimental era formada por quatro segmentos, respeitando-se o padrão de ordem vocabular da língua Karajá, conforme anteriormente mencionado, e descrito de forma mais detalhada no exemplo (16) abaixo:

(16)	<i>Kruku</i>	<i>sohoji-sohoji</i>	<i>òra-tyreki</i>	<i>r-y-i-reri</i>
	PASSARINHO	UM-UM	TRONCO-EM	COP
	sujeito	distributivo	locativo	verbo
	seg1	seg2	seg3	seg4

Aos participantes, cabia a tarefa de observar a imagem e ouvir a sentença a que seria exposto por meio de monitor acoplado simultaneamente a um computador PC e a um aparelho de eletroencefalografia. Ao fim

de cada par <imagem,som>, a tarefa era avaliar se os materiais eram compatíveis entre si ou não. Para isso, os participantes utilizaram um teclado em que havia duas teclas disponíveis destacadas pelas cores verde e vermelho. Caso o participante julgasse que a sentença ouvida era compatível com a imagem anteriormente vista, apertava a tecla verde. Se julgasse que a sentença era incompatível com a imagem vista, apertava a tecla vermelha.

FIGURA 1 – Sequência de telas apresentadas aos participantes durante experimento de EEG



Conforme ilustrado pela figura (1) acima, a apresentação do conjunto de estímulos começava com uma cruz de fixação no centro da tela por 500ms, seguida pela imagem/cenário experimental por 1500ms. Em seguida, os quatro segmentos em áudio eram apresentados. Os segmentos foram controlados de modo a apresentarem um tamanho médio de 1500ms cada. O tempo total entre os segmentos foi de aproximadamente 6000ms. Após ouvida a sentença, havia uma pausa de 250ms, seguida da tela de resposta, que tinha um *time-out* de 2000ms, ou seja, mesmo que o participante não apertasse nenhum dos botões de resposta, após dois segundos um novo conjunto de dados seria apresentado. Ao final de cada sequência experimental, uma tela em branco de 1000ms era apresentada antes de passar para um novo conjunto de dados <imagem, sentença>.

A fim de responder se havia ou não compatibilidade entre figura e sentença, os voluntários foram orientados a apertarem um dos dois botões disponíveis para resposta. Os botões de resposta foram balanceados entre os participantes (direita/esquerda) a fim de evitar efeito de lateralidade,

ou seja, impedir que aqueles que já associassem o *certo* com o lado *direito* se beneficiassem dessa associação e vice-versa. Os tempos de resposta foram registrados pelo *software E-prime (Psychology Software Tools, Inc.)*.

Os sinais de EEG foram gravados continuamente pelo *software PyCorder (Brain Vision LLC)* a partir de um sistema de 64 eletrodos acoplados em uma touca plástica (*Brain Products GmbH*) em acordo com o sistema 10-20 estendido (GOMES, 2014). Os eletrodos foram alocados em acordo com o padrão internacional (*International System Locations*), incluindo cinco eletrodos ao longo da linha média (FPz, Fz, Cz, Pz e Oz) e dezesseis canais lateral/temporal, oito em cada hemisfério do cérebro (FP1/ FP2, F3/ F4, F7/ F8, C3/ C4, T3/ T4, T7/ T8, P3/ P4, e P7/ P8). Além disso, outros quarenta e três canais do sistema 10-20 estendido foram utilizados (AF3/ AF4, F1/ F2, F5/ F6, FC1/ FC2, FC3/ FC4, FC5/ FC6, FT7/ FT8, C1/ C2, C5/ C6, CP1/ CP2, CP3/ CP4, CP5/ CP6, TP7/ TP8, P1/ P2, P5/ P6, P7/ P8, PO3/ PO4, PO5/ PO6, PO7/ PO8, CB1/ CB2 e O1/ O2).

Os eletrodos foram referenciados *on-line* pelo eletrodo *terra*, localizado na região frontal, e posteriormente re-referenciados pela média dos canais mastoides (TP9/TP10 – esquerdo/direito). As impedâncias foram mantidas abaixo de 10 k Ω . O EEG foi amplificado e digitalizado em uma frequência de 1000Hz. Depois de gravados, a frequência de amostragem dos dados foi mantida em 500Hz e foram utilizados os filtros de passa-baixa de 30Hz e de passa-alta de 0.1Hz., metodologia utilizada pelo laboratório ACESIN em acordo com os padrões internacionais de pesquisas na área (TANNER *et al.*, 2016). Posteriormente, foram geradas médias – *averages* – de ERPs para cada sujeito por eletrodo em épocas abrangendo entre 200ms e 1200ms em relação ao *onset* do estímulo alvo. As épocas caracterizadas por piscadas ou movimento muscular excessivo foram automaticamente rejeitadas pelo *Software Vision Analyzer (Brain Products GmbH)*, programa utilizado para filtragem e promediação dos dados gravados.

A extração de ERPs (potenciais bioelétricos relacionados a eventos) é um instrumento de avaliação cognitiva eletromagnético cuja aplicação se estende às pesquisas linguísticas no campo da neurociência da linguagem desde os anos 1980, tendo sido desenvolvido a partir do eletroencefalograma (EEG) inventado pelo psiquiatra alemão Hans Berger

e utilizado em humanos desde 1929.⁸ A técnica de ERP é considerada muito apurada pois registra com uma precisão de milissegundos respostas do sistema nervoso central à estimulação motora ou sensorial sendo, assim, considerado uma técnica de medição *on-line*. O EEG fornece uma amostragem temporal contínua da atividade elétrica do cérebro. Tal amostragem pode estar relacionada a diferentes tipos de fenômenos linguísticos e é monitorada, com auxílio de software específico para este fim, por um aparelho de EEG acoplado a um computador que expõe os participantes do experimento a uma determinada tarefa linguística pré-estabelecida pelo pesquisador linguista. Após a aquisição dos sinais elétricos produzidos no córtex e monitorados pelo EEG, se faz necessário o tratamento (aplicação de filtros) dos mesmos para que possa ser possível analisar os dados obtidos (GESUALDI; FRANÇA, 2011). A filtragem é, portanto, uma parte essencial e indissociável da pesquisa de ERP.

Os ERPs são compostos por uma sequência de ondas caracterizadas por sua latência, amplitude e polaridade. O ERP geralmente apresenta valor instantâneo de 10 a 1000 vezes menor do que o EEG de fundo e por isso não pode ser visualizado. Para que possa haver a visualização é necessário realizar média de várias épocas.⁹ Este procedimento se justifica por assumir-se o EEG espontâneo como um ruído branco gaussiano de média zero e os ERPs como as únicas respostas que são realmente sincronizadas com o estímulo. Deste modo, o efeito da promediação é aumentar a relação sinal/ruído (SNR), assim permitindo a visualização do efeito específico do estímulo, no caso, linguístico. (MARQUES, 2011, p. 30)

Assim, a promediação é um tratamento matemático dos dados utilizado para consolidar a média de um mesmo trecho, em torno de

⁸ Assim como os métodos utilizados na psicolinguística experimental, a técnica de ERP é igualmente não-invasiva permitindo, dessa forma, sua utilização em diversas pesquisas sobre cognição e linguagem em diferentes grupos humanos, de diferentes faixas etárias, com ou sem patologias e sem prejuízo ou desgaste para o participante do experimento. Para realizar a técnica, basta servir-se dos equipamentos e softwares necessários para a aferição, tratamento e estatísticas dos componentes que se deseje averiguar.

⁹ Épocas ou eventos são janelas de trabalho no contínuo temporal do EEG demarcadas para estudo.

um determinado momento acoplado no tempo, para todos os estímulos averiguados. Esse tratamento tem o intuito de excluir ruídos aleatórios concorrentes ao sinal relacionado ao evento estudado, como, por exemplo, aqueles oriundos de movimentos musculares, piscar de olhos, interferência elétrica causada por algum equipamento na sala de testagem etc. Sendo o ruído aleatório, a média obtida após a promediação tenderia a zero. Portanto, ao final desta operação de “limpeza” emergiria apenas o sinal relacionado ao evento linguístico averiguado. Ou seja, uma grande média (ou média abrangente) das janelas de evento vinculadas ao tempo de exibição dos estímulos elaborados para esta investigação (LAGE, 2005).

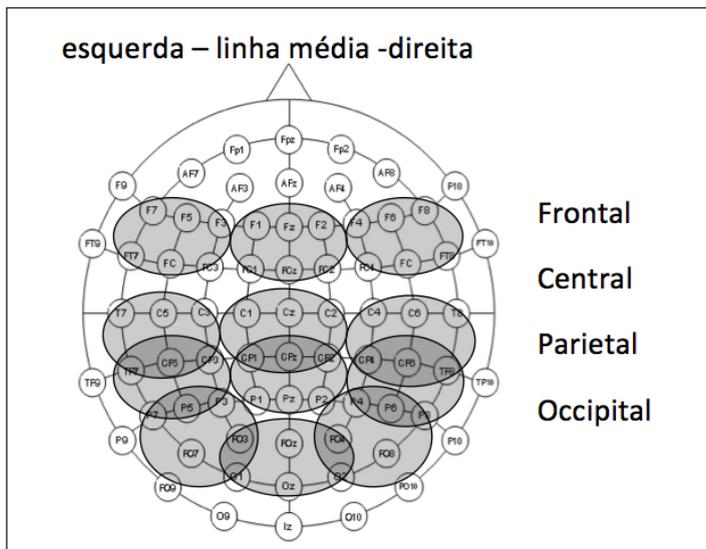
O componente N400 foi descoberto por Kutas e Hillyard no final da década de 70. Os pesquisadores elaboraram um experimento buscando elicitare o componente P3 (já amplamente averiguado na época para reconhecimento de palavras) investigando o papel do contexto no reconhecimento de palavras em sentenças como “*eyebrows*” em “*He shaved off his mustache and eyebrows.*” (Ele barbeou o bigode e a sobrelha dele). Os resultados demonstraram um pequeno componente negativo ao invés de positivo como inicialmente esperado. A fim de observar mais claramente esse efeito, os autores aplicaram um novo experimento com sentenças que continham uma palavra que em vez de ser não esperada pelo contexto era semanticamente anômala. Sentenças como “*He shaved off his mustache and city*” (Ele barbeou o bigode e a cidade dele) e “*I take my coffee with cream and dog*” (Eu tomei meu café com creme e cachorro). Nesse segundo experimento, o efeito de negatividade foi mais explícito uma vez que o mesmo elicitou picos de polaridade negativa bem altos em torno de 400ms no curso do tempo de apresentação da palavra (KUTAS; HILLYARD, 1980). Desde então, este tem sido o componente mais estudado na área da linguagem sendo relacionado principalmente ao conteúdo semântico das palavras per se e às teorias de acesso lexical já que parece fornecer pistas sobre o armazenamento de palavras (FRANÇA *et al.*, 2008). No Brasil, o estudo do componente N400 foi averiguado pioneiramente na tese de doutoramento da professora Aniela Improta França (FRANÇA, 2002) intitulada *Concatenações linguísticas: um estudo de diferentes módulos cognitivos na aquisição e no córtex*. Topograficamente, o N400 tende a apresentar maior ativação na região central e parietal esquerda do cérebro.

Outro componente estudado pela neurociência da linguagem, o P600, foi mais recentemente elicitado por Osterhout e Holcomb (1992) e

desde então tem sido relacionado a anomalias e incongruências sintáticas podendo ser elicitado, por exemplo, em sentenças do tipo **The broker persuaded to sell the stock* (O corretor convenceu a vender o estoque.), em contraste com sentenças bem formadas como *The broker hoped to sell the stock* (O corretor esperava vender o estoque.). Gouvea *et al.* (1995) também encontraram evidências de alta amplitude no componente P600 para sentenças que geram análises temporariamente incorretas (conhecidas na área da psicolinguística como sentenças *garden path*). O P600 também já foi elicitado em anomalias não linguísticas em áreas como música (PATEL *et al.*, 1998) e matemática (MARTÍN-LOECHES *et al.*, 2006). Na área da linguagem, além de violações sentenciais, o componente de P600 parece refletir também violações sintagmáticas como, por exemplo, violações de concordância de número, tempo, gênero e caso, conforme mencionado por Gouvea *et al.* (1995) e investigado por Gomes (2014), Gomes e França (2015) e Gomes e França (2013). A positividade tardia (*Late Positive Component* ou *Late Positive Complex*) já foi elicitada também em estudos sobre quantificação de enunciados (JIANG *et al.*, 2009; KAAAN *et al.*, 2007), o que parece evidenciar que sentenças quantificadas apresentam processamento mais complexo, pois os efeitos de incongruência quantitacional são capturados em momento posterior ao de outros fenômenos da linguagem já averiguados. Kim e Osterhout (2005) avaliam que esse componente pode ser observado em itens cuja incongruência semântica pode levar a uma reanálise sintática.

Para as análises apresentadas a seguir, componentes de interesse N400 e P600 foram inspecionados visualmente e tiveram seus picos registrados automaticamente pelo Vision Analyzer. Para cada um dos canais, foram capturados o pico de amplitude e a voltagem dos componentes de interesse em todos os segmentos e condições. O componente N400 foi inspecionado dentro de uma janela de 350-550ms (capturando ampla negatividade) e o componente P600 foi inspecionado dentro de uma janela de 550-800ms (capturando ampla positividade) após o *onset* do estímulo alvo. Grandes médias – *Grand-averages* – foram formadas pela promediação das médias dos ERPs de interesse de cada participante.

FIGURA 2 – Regiões de interesse baseadas em proximidade anatômica



As medidas dependentes, isto é, a amplitude e latência dos componentes N400 e P600 foram analisadas a partir de análise de variância (ANOVA) e testes paramétricos (Teste-T). As ANOVAS foram aplicadas entre os fatores *Condição* (distributiva, distributiva parcial, coletiva, extra-elemento, não exaustiva e unicidade), *Lateralidade* (hemisfério esquerdo, direito e linha média) e *Anterioridade* (região frontal, central, parietal e occipital). Os testes paramétricos foram realizados para averiguação de diferença significativa entre a condição distributiva padrão e as demais condições experimentais. Portanto, realizamos quatro análises distintas a partir dos dados de ERPs obtidos:

- i. ANOVA das amplitudes dos picos das ondas para o componente N400 por condição, lateralidade e anterioridade. Teste-t contrastando as médias obtidas com as da condição padrão distributiva.
- ii. ANOVA das amplitudes dos picos das ondas para o componente P600 por condição, lateralidade e anterioridade. Teste-t contrastando as médias obtidas com as da condição padrão distributiva.

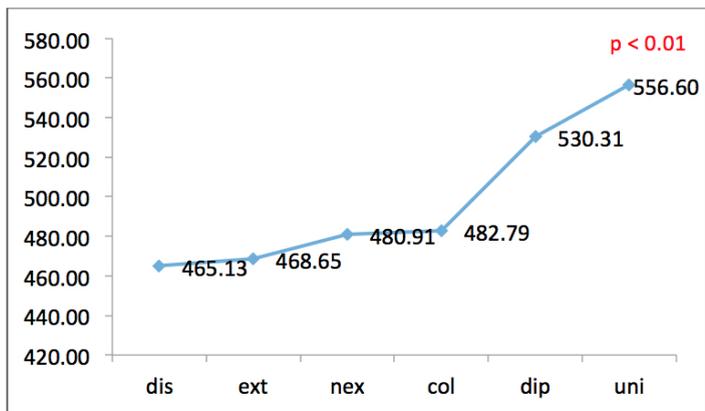
A estatísticas baseadas nos componentes N400 e P600 são as medidas mais discutidas na literatura de neurociência da linguagem

sendo, assim, o melhor ponto de partida para um experimento desta natureza. Além disso, trabalhos que averiguaram diferentes aspectos da quantificação de enunciados elicitaram N400 em sentenças quantificadas, evidenciando um custo de processamento inicial maior para enunciados que envolvem procedimentos como contagem e individuação (CHIARELLI *et al.*, 2011). Também verificamos efeitos de positividade tardia, interpretada por Jiang, Tan e Zhou (2009) como um esforço de integração de elementos quantificadores em sentenças que apresentem inadequação quantificacional. Portanto, para as análises apresentadas a seguir, extraímos a amplitude dos picos de N400 e P600 em todas as condições experimentais a partir das grandes médias dos segmentos 2 e 3 das sentenças manipuladas.

4.3 Dados Comportamentais

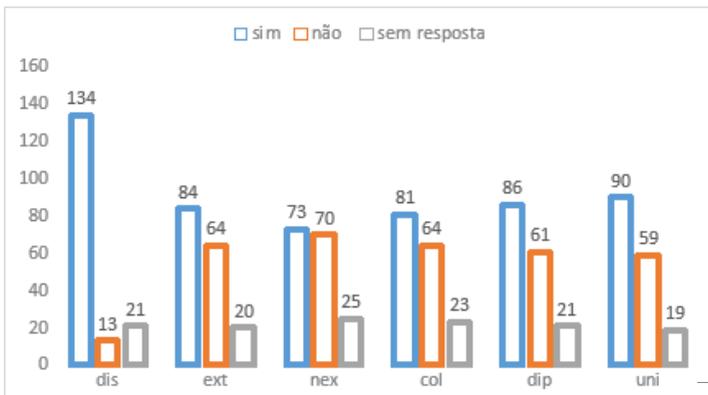
A análise realizada por meio de test-t dos tempos de reação (RT – em milésimos de segundos) obtidos dos participantes nos permitiu observar quais condições apresentaram maior tempo médio de decisão em relação à condição distributiva padrão. Os resultados obtidos apontam que apenas a condição **unicidade** ($t = -2.585$, $df = 21.182$, $p\text{-value} = 0.0172$) elicitou RTs significativamente maiores que a condição distributiva, o que sugere maior custo cognitivo para a decisão dessa condição. Esta análise dos RTs brutos nos mostra que o cenário unicidade foi o que causou maior estranhamento durante a tarefa de pareamento com uma sentença que continha o numeral distributivo *sohoji-sohoji*. As outras condições não apresentaram diferenças significativas, conforme pode ser conferido no gráfico (1) abaixo.

GRÁFICO 1 – RT bruto por condição experimental



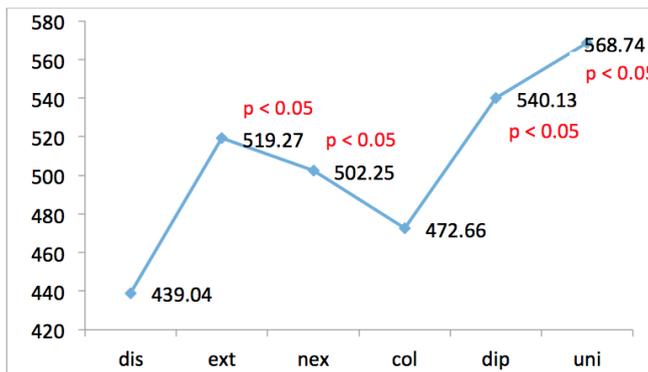
De uma forma geral, observa-se uma tendência em aceitar todas as condições como cenários compatíveis para sentenças com *sohoji-sohoji*, conforme o gráfico 2 abaixo. Entretanto, após a realização de testes de chi-quadrado para cada condição, observou-se que apenas as condições distributiva ($X\text{-squared} = 99.5986$, $df = NA$, $p\text{-value} = 0.0004998$) e unicidade ($X\text{-squared} = 6.4497$, $df = NA$, $p\text{-value} = 0.02249$) apresentaram uma diferença significativa entre as respostas “sim” e “não”. O fato da condição distributiva apresentar o maior índice de aceitação confirma que este é o melhor cenário para as sentenças experimentais averiguadas. A aceitação da condição unidade parece ser resultado de uma tentativa de leitura cardinal do numeral reduplicado, pois como veremos nos dados bioelétricos, essa condição foi a que obteve maior custo cognitivo durante o processamento. Embora as outras condições não tenham capturado um efeito significativo neste teste ($p > 0.05$), observamos que medidas mais reflexas, como os RTs e os potenciais bioelétricos parecem revelar que há diferenças de processamento entre as diferentes condições averiguadas e a condição distributiva padrão. O *time-out* programado para 2000ms gerou alguns itens “sem resposta” que foram excluídos das análises.

GRÁFICO 2 – Índices de resposta por condição



Como em todas as condições obtivemos maiores índices para as respostas “sim”, aplicamos novamente o teste t entre condições para verificar os RTs apenas dos cenários aceitos pelos participantes, excluindo os cenários julgados como não compatíveis. As condições **distributiva parcial** e **unicidade** mantiveram-se com as médias mais altas, neste caso ambas apresentaram diferenças significativas. Entretanto, as condições **não exaustiva** e **extra elemento** também obtiveram tempos significativamente mais altos. A condição coletiva não apresentou diferença de RT significativa quando aceita pelo falante, neste caso a média foi estatisticamente semelhante a da condição distributiva padrão. Os p-valores e médias podem ser conferidos no gráfico 3 abaixo.

GRÁFICO 3 – RT dos pareamentos aceitos por condição experimental



Esses resultados evidenciam a importância dos dados cronométricos para nossa análise, pois os índices de aceitabilidade não mostraram ser um fator decisivo dentre todos os cenários apresentados. Além disso, embora a condição desviante unicidade tenha sido aceita como cenário válido para a sentença experimental, a análise dos RTs evidenciam que essa aceitabilidade foi efetuada em um tempo significativamente maior que para a condição distributiva padrão. Sem a análise dos tempos de decisão em milissegundos também seria impossível capturar a distinção significativa entre as demais condições experimentais observadas, já que a aceitabilidade se manteve constante para todos os cenários. Entretanto, a partir dos RTs ainda não conseguimos obter evidências de que haja diferença de processamento entre as condições não-exaustiva e extra-elemento, nem capturar se há esforço cognitivo maior para aceitar a condição, como esperado.

Os resultados comportamentais apresentados nesta seção foram capturados após o processamento das sentenças, ou seja, de maneira *off-line*. Para análise de dados *on-line*, capturamos também as respostas eletrofisiológicas obtidas através da técnica de extração de ERPs. Esta técnica teve como finalidade observar o processamento das sentenças experimentais de momento a momento durante o curso temporal de compreensão e pareamento das sentenças. Os resultados obtidos a partir dessa metodologia serão apresentados na próxima subseção e apontarão novas reflexões no que concerne a compreensão do processamento do distributivo *sohoji-sohoji* e revelarão dados interessantes sobre como se dá a aceitabilidade dos diferentes cenários averiguados.

4.4 Dados de ERP

Para a análise estatística, utilizamos primeiro o modelo three-way ANOVA (*within subjects*) que incluiu os fatores condição, lateralidade e ROI, com correção dos valores apresentados (*Greenhouse-Geisser*). Após tratamento dos dados elicitados, observamos que o componente N400 apresentou efeito principal significativo para o fator *condição* ($F(5,25) = 7.56 p < 0.0000003$) já no segundo segmento das sentenças experimentais, correspondente ao numeral reduplicado *sohoji-sohoji*. Por outro lado, o componente P600 elicitou efeito principal para o fator *condição* ($F(5,25) = 14.7 p < 0.0017565$) apenas no segmento 3, correspondente ao sintagma posicionado (PP) locativo. Não foram encontrados efeitos principais

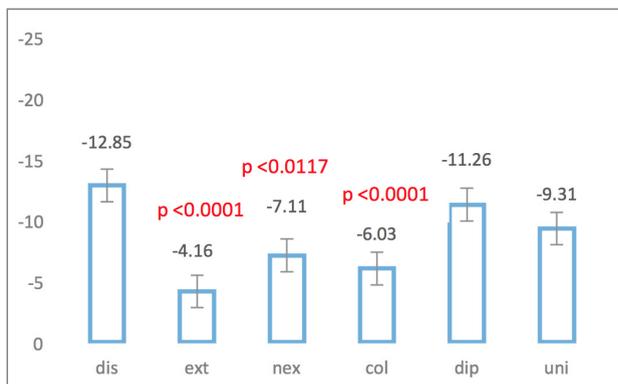
para os fatores ROI ($F(11,55) = 1.19$ $p < 0.3472345$) ou lateralidade ($F(2,10) = 1.90$ $p < 0.2238272$) no componente P600 nem para os fatores ROI ($F(11,55) = 1.41$ $p < 0.195080$ $p < 0.2820406$) ou lateralidade ($F(2,10) = 2.55$ $p < 0.1680510$) no componente N400. Tampouco encontramos interação entre os fatores mencionados. Assim, parece que apenas o fator condição parece influenciar, de fato, as análises entre os resultados obtidos. Por isso, optamos por realizar uma análise incluindo todos as ROIs para esse artigo. Futuramente, outras análises poderão ser realizadas para confirmar as tendências aqui apresentadas.

Posteriormente, foi aplicado um teste t que contrastou a condição distributiva padrão com as outras condições experimentais. Abaixo apresentamos os resultados encontrados para as medidas de amplitude¹⁰ (μV) dos dados de ERPs das Grandes Médias para os componentes N400 e P600.

4.4.1 Componente N400 – negatividade após *onset* de *sohoji-sohoji*

O componente N400, averiguado após o início do segmento que continha o numeral reduplicado *sohoji-sohoji*, apresentou resultados significativos ao comparar a condição distributiva padrão ($-13 \mu V$) com a condição **não exaustiva** ($-7 \mu V$, $t(120)=2.56$ $p < 0.0117$), **coletiva** ($-6 \mu V$, $t(120)=4.88$ $p < 0.0001$) e **extra elemento** ($-4 \mu V$, $t(120)=6.13$ $p < 0.0001$). As condições unicidade e distributiva parcial não apresentaram diferenças significativas.

GRÁFICO 4 – Médias de amplitude do componente N400 por condição

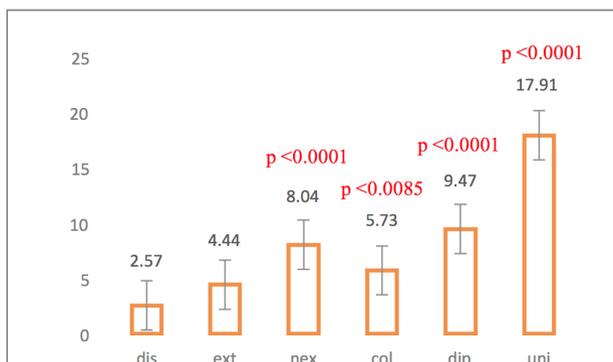


¹⁰ Média dos picos de amplitude dos componentes em microvolts por condição.

4.4.2 Componente P600 – positividade tardia após o *onset* do PP locativo

O componente P600, averiguado após o início do segmento que continha o PP locativo da eventualidade, elicitou diferença significativa entre as condições **unicidade** (18 μ V, $t(118)=17.75$ $p < 0.0001$), **distributiva parcial** (9 μ V, $t(118)=4.91$ $p < 0.0001$), **não exaustiva** (8 μ V, $t(118)=4.17$ $p < 0.0001$) e **coletiva** (6 μ V, $t(118)=2.67$ $p < 0.0085$) ao compararmos com a condição distributiva padrão (3 μ V).

GRÁFICO 5 – Médias de amplitude do componente P600 por condição



4.5 Discussão dos resultados e análises

Este experimento teve como fio condutor a busca por uma compreensão microscópica do processamento do distributivo *sohoji-sohoji*. Para este experimento utilizamos contextos previamente elaborados que foram fornecidos por meio de imagens controladas por *design* específico, como descrito no início desta seção. A metodologia utilizada possibilitou as análises dos dados eletrofisiológicos capturados pelo EEG durante o processamento das sentenças, medida *on-line*, que refletiriam o custo cognitivo durante o processamento das sentenças ao serem pareadas com diferentes cenários quantitativos. Os dados foram então filtrados para a extração e análise das amplitudes (μ V) capturadas a partir da delimitação dos componentes N400 (janela de tempo de 250-500 ms) e P600 (janela de tempo de 550-800 ms). Os resultados obtidos apontam que a aceitabilidade de diferentes condições experimentais em dados comportamentais não reflete o custo cognitivo para o pareamento entre os diferentes cenários apresentados e o numeral reduplicado

sohoji-sohoji, uma vez que os participantes tenderam a aceitar todas as condições. Porém, houve uma tendência de tempos de reação (RTs) mais lentos para cenários desviantes do padrão distributivo, a saber, aqueles que permitiam leituras distributivas intermediárias e não pluralizadas. Portanto, os resultados parecem confirmar a hipótese previamente levantada de que o processamento de leituras cumulativas, apesar de possível, seria mais custoso do que sentenças não cumulativas quando há a presença de um elemento que delimite o escopo quantificacional, neste caso o numeral reduplicado *sohoji-sohoji*. Como discutido abaixo, os dados *on-line* sugerem que as propriedades de cumulatividade, distributividade e cardinalidade exercem efeitos durante o processamento de sentenças com *sohoji-sohoji* em Karajá. Essas evidências parecem dialogar com teorias vigentes acerca do processamento quantificacional e da semântica de numerais distributivos já observados em outras línguas naturais.

No que diz respeito aos dados de amplitude dos ERPs, observou-se que o componente N400 elicitou amplitudes mais baixas para as condições **coletiva** e **extra-elemento** no segmento 2, correspondente ao numeral reduplicado *sohoji-sohoji*. Já o componente P600 elicitou amplitudes mais altas para as condições **não exaustiva**, **coletiva**, **distributiva parcial** e **unicidade** no segmento 3, correspondente ao PP locativo. Esses resultados parecem refletir o custo cognitivo de processamento do numeral reduplicado frente aos diferentes cenários apresentados durante o experimento. O componente N400 revela que o cenário distributivo prototípico, aquele que temos um indivíduo para cada eventualidade, apresenta a maior média de amplitude, seguido pela condição distributiva parcial. Tal evidência parece indicar que o processamento de *sohoji-sohoji* engajaria o sistema cognitivo dos indivíduos em uma operação distributiva, que exige maior custo de processamento em nível reflexo. Já para a medida P600, registrada no segmento locativo que segue *sohoji-sohoji*, observamos que as condições que infringem a leitura do tipo evento-distributivo são aquelas que apresentam maiores amplitudes, revelando que é após o início do procedimento de pareamento do elemento contido na eventualidade que emergem medidas que revelariam quais propriedades seriam mais facilmente processadas em sentenças com numeral distributivo.

A condição **unicidade** foi a que elicitou médias mais altas tanto no componente P600 (18 μ V) quanto nos dados comportamentais brutos

(557 ms) e filtrados por aceitabilidade (568 ms). O que sugere que o numeral reduplicado realmente exige a pluralização do enunciado, no que diz respeito as relações estabelecidas pelos indivíduos e eventualidades disponíveis, em acordo com o que é esperado por elementos distributivos.

A condição **distributiva parcial** também apresentou uma média de amplitude alta para o componente P600 (9 μ V) e também foi a segunda medida mais alta dos tempos de reação filtrados por aceitabilidade (540 ms). Estes resultados indicam que leituras intermediárias, apesar de aceitas, apresentam grande custo de processamento. Neste caso, como *sohoji-sohoji* é formado a partir da reduplicação do numeral *um* também podemos analisar estes resultados como provenientes da falha de cardinalidade numeral apresentada pelo cenário apresentado, ou seja, um problema na operação de contagem.

A condição **não exaustiva** elicitou o componente P600 (8 μ V) e tempos de reação filtrados por aceitabilidade (502 ms) significativos. Nesta condição, um elemento da eventualidade não estava em uma relação R com nenhum indivíduo do cenário. O custo elevado de processamento parece evidenciar que a interpretação de sentenças com *sohoji-sohoji* exige que os subeventos do cenários opere preferencialmente sob o escopo de um quantificador universal.

A condição **coletiva** apesar de elicitar o componente P600 (6 μ V) apresentou médias mais baixas que as outras condições acima mencionadas. Além disso, não foi capturado nenhum efeito nos dados comportamentais. Acreditamos que esse efeito obtido nesta condição poderia ter sido potencializado se fossem averiguados eventos com verbos transitivos que permitissem uma leitura coletiva do evento e não apenas do locativo. Essa hipótese é suportada pelo fato de que nas sentenças averiguadas os participantes aceitaram sem maiores custos cognitivos o cenário coletivo apresentado após o processamento completo da sentença, uma vez que não encontramos efeitos significativos nos RTs.

A condição **extra elemento** não obteve diferenças significativas no componente de P600, ou seja, um cenário em que um indivíduo não participe do evento em distribuição não afetou a amplitude positiva averiguada (4 μ V). Entretanto, nos resultados comportamentais filtrados por acurácia (519 ms) esta condição apresentou resultados significativos. Esses resultados indicam que em um nível mais reflexo, *on-line*, a universalidade sobre eventos é menos custosa do que a universalidade

sobre indivíduos durante a operação de distributividade de *sohoji-sohoji*. A diferença encontrada no RT pode ter sido efeito de uma reinterpretação pós-processamento sentencial.

Em relação aos resultados significativos encontrados no componente N400 para as condições **extra elemento** ($-4 \mu\text{V}$), **não exaustiva** ($-7 \mu\text{V}$) e **coletiva** ($-6 \mu\text{V}$), uma proposta é a de que nestas condições há uma demora maior para o engajamento do sistema de memória em computar a distributividade, pois médias mais baixas refletem menor esforço cognitivo. É, portanto, adequado propor que esses resultados refletem que as incompatibilidades semânticas de universalidade e diferenciação presentes nesses cenários levaria o sistema cognitivo a inicialmente não se engajar na verificação de compatibilidade de tais cenários, procrastinando a verificação para um momento posterior do processamento, conforme verificado pelos resultados obtidos no componente P600 e nos dados comportamentais. A negatividade elicitada pelo componente N400 para as condições **distributiva padrão** ($-13 \mu\text{V}$) e **distributiva parcial** ($-11 \mu\text{V}$) parece ser resultante do efeito de incrementacionalidade dos quantificadores, que seriam interpretados em estágios, assim como acontece com a negação (URBACH; KUTAS, 2010; e URBACH; DELONG; KUTAS, 2015). Logo, a resposta neurofisiológica obtida por medidas de amplitude mais altas para condições distributivas do teste, em um momento inicial do processamento da sentença, refletiria que o participante está engajado no procedimento de uma operação de maior custo cognitivo, no caso, o procedimento da verificação de distributividade numeral. Já a negatividade elicitada para a condição **unicidade** ($-9 \mu\text{V}$) pode ser consequência da implausabilidade semântica em se comparar um cenário não pluralizado com uma sentença distributiva, o que geraria uma resposta neurofisiológica do participante que refletiria o susto em processar uma sentença operada por um numeral distributivo após ser exposto a cenário singular, que deveria ser imediatamente descartado.

5. Considerações finais

Este trabalho procurou demonstrar como a disponibilidade de técnicas cada vez mais sensíveis para averiguar como o estudo do processamento da linguagem pode enriquecer as pesquisas de campo tradicionais e análises formais da linguagem. Nesta perspectiva

transdisciplinar, o experimento de EEG para a extração ERPs teve como objetivo averiguar como os participantes processariam sentenças distributivas com o numeral reduplicado *sohoji-sohoji* quando expostos a cenários que apresentavam leituras intermediárias, não prototipicamente distributivas.

Nossos resultados demonstram que os componentes N400 e P600 foram elicitados quando houve incompatibilidades entre cenário e sentença. A reanálise de condições que violavam a pluralidade, cardinalidade, diferenciação ou exaustividade, foi capturada por uma ampla positividade no componente P600, elicitada a partir dos elementos que ocorriam como sintagmas locativos dos eventos. Em um momento anterior, capturamos um efeito de ampla negatividade no componente N400 em condições que engajavam os participantes em um procedimento quantitativo (distributiva padrão e distributiva parcial) ou que descartavam de imediato a plausibilidade entre o cenário investigado e a sentença distributiva (unicidade).

Esses resultados complementam aqueles encontrados nas medidas cronométricas *off-line* que apontam que *sohoji-sohoji* poderia ou atuar como um quantificador ambíguo, permitindo que mais de uma leitura seja licenciada para a sentença (leituras evento-distributiva ou participante-distributiva) ou licenciaria diferentes cenários de verificação para a mesma leitura, em uma perspectiva próxima àquela proposta por Cable (2014) que propõe uma denotação semântica capaz de dar conta dos diferentes cenários de verificação inerentes aos quantificadores distributivos em Tiglit cuja resolução dependerá, sobretudo, do contexto fornecido.

Além disso, se levarmos em consideração a proposta de Choe (1987), de que o operador distributivo é um quantificador universal que tem um conjunto em restrição e um conjunto em escopo, podemos considerar que o numeral reduplicado *sohoji-sohoji* como *default* colocaria a eventualidade no conjunto em restrição, ou seja, a leitura evento-distributiva seria menos custosa. Isto porque é aos membros do conjunto de eventualidades que o operador universal (\forall) deveria ser aplicado. Ou seja, a não exaustividade dos elementos da eventualidade da sentença seria a leitura que apresentaria maior custo cognitivo porque necessita de reanálise, conforme foi constatado nos dados neurofisiológicos.

A distributividade é um conhecimento que faz parte da cognição humana e é encontrada em diferentes línguas naturais por meio de diversos mecanismos morfológicos e sintáticos. Esperamos que este trabalho possa dar visibilidade para o fenômeno de distributividade numeral encontrado nas línguas ameríndias que podem conter propriedades relevantes para o estudo das áreas pertinentes à linguística e, por conseguinte, para o conhecimento da faculdade da linguagem humana.

Referências

CABLE, S. Distributive numerals and distance distributivity in tlingit (and beyond). *Language*, Washington, Linguistic Society of America, v. 90, n. 4, p. 562-606, 2014.

CHIARELLI, V.; YAGOUBI, R.; MONDINI, S.; BISIACCHI, P.; SEMENZA, C. The syntactic and semantic processing of mass and count nouns: an ERP study. *PLoS ONE*, San Francisco, CA, v. 6, n. 10, p. 1-15, 2011.

DOWTY, D. Collective Predicates, Distributive Predicates, and All. In: ESCOL, 3rd, 1987, Columbus. *Proceedings...* Columbus: Ohio State University, 1987.

FRANÇA, A. I. *Concatenações linguísticas: estudo de diferentes módulos cognitivos na aquisição e no córtex*. 2002. Tese (Doutorado em Linguística) – UFRJ, Rio de Janeiro, 2002.

FRANÇA, A. I.; LEMLE, M.; GESUALDI, A.; CAGY, M.; INFANTOSI, A. F. C. A neurofisiologia do acesso lexical: palavras em português. *Revista Veredas*, v. 2, p. 34-49, UFJF, Juiz de Fora, 2008.

GESUALDI, A. da R.; FRANÇA, A. I. Event-related brain potentials (ERP): an overview. *Revista Linguística*, UFRJ, Rio de Janeiro, v. 7, n. 2, p. 24-42, 2011.

GIL, D. *Distributive numerals*. 1982. Dissertation (Ph.D.) – University of California, Los Angeles, 1982.

GOMES, J. N. *Investigating the dissociation between n400 and p600 effects on the syntax semantics interface: an ERP study*. 2014. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2014.

GOMES, J.; FRANÇA, A. I. A técnica de ERP: investigando a assimetria sujeito-objeto na interface sintaxe-semântica com EEG. *Letras de Hoje (Online)*, Pontifícia Universidade Católica do Rio Grande do Sul, v. 4, p. 152-163, 2015.

GOMES, J. N.; FRANÇA, A. I. The ample semantic scope of minute language computations: an ERP study of words in Portuguese. *Revista da ANPOLL*, Anpoll, v. 34, p. 309-346, 2013.

GOUVEA, A. C.; PHILLIPS, C.; KAZANINA, N.; POEPPPEL, D. The linguistic processes underlying the P600. *Language and Cognitive Processes*, Taylor & Francis Online, v. 25, p. 149-188, 1995.

HEIM, I. *The semantics of definite and indefinite noun phrases*. 1982. Thesis (PhD) – GLSA, University of Massachusetts, Amherst, 1982.

JIANG, X.; TAN, T.; ZHOU, X. Processing the universal quantifier during sentence comprehension: ERP evidence. *Neuropsychologia*, Elsevier, v. 47, p. 1799-1851, 2009.

KAAN, E.; DALLAS, A. C.; BARKLEY, C. M. Processing bare quantifiers in discourse. *Brain Research*, Elsevier, v. 1146, p. 199-209, 2007.

KIM, A.; OSTERHOUT, L. The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, Elsevier, v. 52, p. 205-225 2005.

KRATZER, A. On the plurality of verbs. In: DÖLLING, J.; HEYDE-ZYBATOW, T. (Ed.). *Event structures in linguistic form and interpretation*. Berlin: Mouton de Gruyter, 2007.

KRATZER, A. The event argument and the semantics of verbs. 2003. Manuscrito. Disponível em: <<http://semanticsarchive.net>>.

KRIFKA, M. The origins of telicity. In: ROTHSTEIN, S. (Ed.). *Events and grammar*. Dordrecht: Kluwer, 1998. p. 197-235.

KUTAS, M.; HILLYARD, S. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, American Association for the Advancement of Science, v. 207, p. 203-205, Jan. 1980.

LAGE, A. C. *Aspectos neurofisiológicos de concatenação e idiomacidade em português do Brasil: um estudo de potenciais bioelétricos relacionados a eventos linguísticos (ERPs)*. 2005. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

- LYONS, J. *Semantics*. London: Cambridge University Press, 1977. 2 v.
- MARQUES, F. *Eletrofisiologia da idiomaticidade em indivíduos com Síndrome de Asperger: um estudo de ERPs*. 2011. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.
- MARTÍN-LOECHES, M.; CASADO, P.; GONZALO, R., DE HERAS, L.; FERNÁNDEZ-FRÍAS, C. Brain potentials to mathematical syntax problems. *Psychophysiology*, Wiley Online Library, v. 43, p. 579-591, 2006.
- MAY, R. *The grammar of quantification*. 1977. Dissertation (Ph.D.) – MIT, 1977.
- MAY, R. *Logical form: its structure and derivation*. Cambridge, Massachusetts: MIT Press, 1985.
- MULLER, A. Distributividade: o caso dos numerais reduplicados em Karitiana. *Cadernos de Estudos Linguísticos*, Campinas, v. 54, n. 2, p. 223-243, 2012.
- OLDFIELD, R. C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, Elsevier, v. 9, p. 97-113, 1971.
- OSTERHOUT, L.; HOLCOMB, P. J. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, Elsevier, v. 31, n. 6, p. 785-806, 1992.
- PATEL, A. D.; GIBSON, E.; RATNER, J.; BESSON, M.; HOLCOMB, P. J. Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, MIT Press, v. 10, n. 6, p. 717-733, 1998.
- SILVA, C. O. da. *O processamento do numeral distributivo sohoji-sohoji na língua karajá*. 2016. Tese (Doutorado em Linguística) – Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.
- TANNER, D.; NORTON, J. J.; MORGAN-SHORT, K.; LUCK, S. J. On high-pass filter artifacts(they're real) and baseline correction (it's a good idea) in ERP/ERMF analysis. *Journal of Neuroscience Methods*, Elsevier, v. 15, n. 266, p. 166-170, Jun. 2016.
- TUNSTALL, S. *The interpretation of quantifiers: semantic and processing*. 1998. Dissertation (PhD) – University of Massachusetts, Amherst, 1998.

URBACH, T. P.; DELONG, K.; KUTAS, M. Quantifiers are incrementally interpreted in context, more than less. *Journal of Memory and Language*, Elsevier, v. 83, p. 79-96, 2015.

URBACH, T. P.; KUTAS, M. Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. *Journal of Memory and Language*, Elsevier, v. 63, p. 158-179, 2010.

The processing of the reduplicated numeral *sohoji-sohoji* in Karajá: An ERP investigation during sentence comprehension

O processamento do numeral reduplicado sohoji-sohoji em Karajá: uma averiguação de ERP durante a compreensão de sentenças¹

Cristiane Oliveira da Silva

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
cristianeolivers@gmail.com

Resumo: Este trabalho tem como objetivo averiguar o processamento do numeral distributivo *sohoji-sohoji*, formado pela reduplicação do numeral *um* em Karajá, sob uma perspectiva transdisciplinar aproximando os campos da linguística teórica e da neurobiologia. O experimento de neurociência da linguagem contribui para elucidar a complexidade do processamento de sentenças com *sohoji-sohoji* quando pareado com diferentes cenários quantitativos. Utiliza-se a técnica de extração de ERP (potenciais bioelétricos relacionados a eventos) para observação e extração dos dados de latência visando a análise estatística dos dados cronométricos monitorados *online* por um eletroencefalograma (EEG). Os resultados das análises de variância e teste *t* demonstraram que a leitura evento-distributiva é a *default* porque a violação de exaustividade nos cenários testados apresentou maior complexidade de processamento do que a violação de exaustividade de participantes. Pluralidade, cardinalidade e diferenciação de eventos também apresentaram efeitos significativos durante a compreensão de sentenças com *sohoji-sohoji*.

Palavras chave: Karajá; distributividade; ERP; processamento quantificacional.

¹ Work based on the Doctoral Dissertation defended at UFRJ in 2016; Advisor: Marcus Maia and Coadvisor: Aniela Improta França.

Abstract: The aim of this paper is to investigate the processing of the reduplicated numeral *sohoji-sohoji* formed by the reduplication of the numeral *one* in Karajá, within an interdisciplinary approach, bringing together formal linguistics and neurobiology. The neuroscience of the conducted language experiment sheds new light on the quantification processing complexity of *sohoji-sohoji* when different quantitative scenarios were previously presented. We use the Event-Related Brain Potential (ERP) extraction technique to observe and extract the wave's amplitude in order to conduct a statistical analysis from the chronometric data monitored online by the electroencephalogram (EEG). The results from statistical analyses (ANOVA and *t* test) conducted in this study suggest that the event-distributive interpretation is the first to be activated, because the violation of event item exhaustiveness was more complex than was the violation of subject exhaustiveness. Plurality, cardinality, and event differentiation also displayed significant effects during Karajá distributive sentence comprehension.

Keywords: Karajá; distributive property; ERP; quantification processing.

Received on January 12, 2017

Approved on May 8, 2017

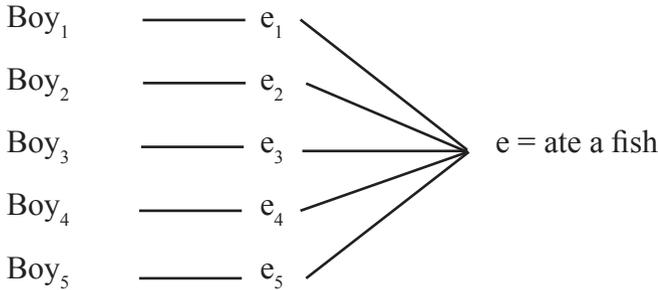
1. Introduction

The quantification of events or individuals is a common operation in natural languages. However, each linguistic system has distinct mechanisms for expressing notions related to quantity. The study of quantifiers in natural languages is a topic of interest in different areas of knowledge, as philosophy and logic, for example, having a range of classical works that attempt to ascertain the functioning of linguistic elements to express quantity. In this paper, we present a processing study that used the Event-Related Brain Potentials (ERP) technique to investigate neurobiological responses during the processing of distributive sentences with the *sohoji-sohoji* quantifier, a reduplicated numeral in the Karajá language.

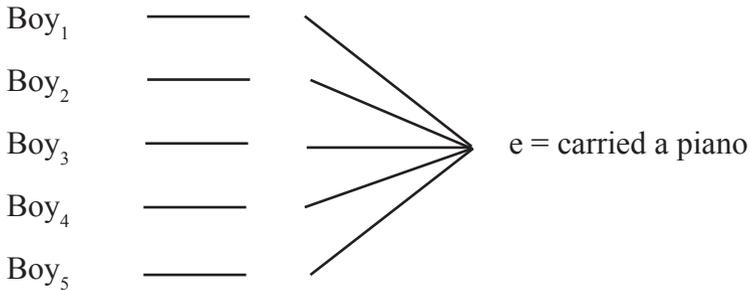
Distributive quantifiers are responsible for a mapping operation between the subevents and individuals of a statement. Therefore, a

prototypical distributive reading should operate on individuals and subevents under the condition of differentiation between mapped individuals and subevents. As shown in example (1), for each individual x such that x is a *boy*, there will be a different sub-event of *eat a fish*.

- (1) Each boy ate a fish.



- (2) The boys carried a piano.



However, delimiting certain types of events does not seem to be a trivial task and non-differentiation of the subevents could then result in a collective reading, as in example (2). Tunstall (1998) also emphasizes the existence of a gradient between distributive and collective readings, i.e., there would be a continuous scale between prototypical distributive and collective readings in which intermediate readings would be available, whether partially distributive or partially collective. Kratzer (2003, 2007) treats these intermediate readings as cumulative.

Krifka (1998) and Kratzer (2003, 2007) argue in favor of the universal cumulative property of lexical items, although this does not guarantee that the lexicon is cumulative in the language syntax. In this light, assuming that lexical items are also cumulative in Karajá syntax, a cumulative reading would always be available in that language. Hence, we could speculate that the processing of a cumulative reading would present a lower processing cost, except when there is an element that delimits the quantifiable scope of the sentence, as, for example, a distributive operator. Therefore, by using the *sohoji-sohoji* reduplicated numeral, a distributive reading for the statement would be forced, even if other readings could be accommodated. In this bias, two questions can be raised in relation to the Karajá language: (i) Would it be possible to process intermediate readings in the presence of the reduplicated numeral? (ii) Would pairing intermediate readings of sentences with *sohoji-sohoji* quantifier present a higher processing cost? We confirmed in a pilot study based on scenario descriptions that scenarios containing intermediate readings could be described by the participants using the *reduplicated* numeral. This fact led to the development of an experiment that would make it possible to verify whether there would be sentence processing differences between the different readings available when paired with the reduplicated numeral in Karajá.

Thus, an ERP experiment was performed to investigate the processing of the *sohoji-sohoji*. The experiment was based on a picture matching design. We attempted to assess the course of sentence processing presented by capturing the reaction times and the neurobiological responses during the experiment. To this end, we ascertained the processing of sentences containing the reduplicated numeral after showing the distributive, collective, and partially distributive scenarios. As shown below, the properties of cumulativity, distributivity and cardinality exert effects during the processing and interpretation of sentences with *sohoji-sohoji* in Karajá. Such evidence seems to run in line with current theories about quantification processing and the semantics of distributional numerals, which have been observed in other natural languages.

2. The reduplicated numeral in Karajá (and other languages)

The Karajá language (Karajá family, Macro-Je stock) presents numeral reduplication to express the distributive property, as can be seen in example (2) below. Numeral reduplication brings to light two types of distributive readings, according to the interpretations presented in (3).

- (2) *weryry* *sohoji-sohoji* *bola-di* *r-e-hu-ra*
 BOY ONE-ONE BALL-POSP 3A-VT-TO.THROW-PST.REC
 Lit: ‘Boy threw the ball one by one’

- (3) a. ‘The boys threw a ball one at a time’ event-distributive
 b. ‘Each boy threw a ball’ participant-distributive

While the event-distributive reading (3a) implies that each individual has thrown the ball in distinct periods of time, the participant-distributive reading (3b) does not present this requirement for the sentence interpretation. In turn, a collective reading does not seem to be allowed, i.e., a reading in which all the boys have thrown a single ball at the same time.

The difference between the readings above seems to come from the semantic ambiguity triggered by the reduplicated numeral, *sohoji-sohoji*. One of the objectives of this paper is to determine whether there would be a preference for some of the readings presented in (3) and whether a collective reading would be immediately discarded, given its incompatibility with the semantics of the distributive numeral, during the processing of distributive sentences in Karajá.

The possibility of a participant-distributive reading requires that the distributive operate over individuals. The impossibility of combining the reduplicated numeral with mass nouns in Karajá, which do not allow the individuation of entities because it is a substance, can be confirmed by impossibility of example (4) below. Sentence (4) demonstrates that the substance *bèè* (water) is incompatible with the distributive operation carried out by *sohoji-sohoji*.

- (4) *Tii bèè sohoji-sohoji-my r-e-wy-reri.
 3P water one-one-POSP 3P-VT-to.transport-PRS.CONT
 He carries the water one at a time/ He carries each water

Numeral reduplication as a distributive operation strategy is relatively common in several languages. According to the WALS² mapping, 85 from 251 languages investigated, present a numeral reduplication strategy to denote adnominal distributive property. Gil (1982) also shows that all languages have some resource to express adverbial distributive property by numeral morphology. It should be noted, however, that these studies cover very few indigenous languages and the Karajá language was not investigated by these papers.

Karitiana (Arikém family) is another indigenous language that has reduplicated numerals to express distributive property. As in Karajá, reduplicated numerals in Karitiana generate different possibilities of interpretation for the same sentence as exemplified below in (5). Muller (2012) assumes, based on syntactic tests, that distributive numerals in Karitiana are adverbial distributive operators. For the author, the two available readings are the result of different modes of individuation that is possible for the subevents in the verbal predicate denotation.

- (5) Õwã nakakot sypomp.sypomp opokakosypi
 BOY BROKE TWO.TWO EGG
 ‘Each boy broke two eggs’ – participant-distributive
 ‘The boys broke two eggs at a time’ – event-distributive

Muller (2012, p. 235)

However, in analyzing the numeral morphology in Tlingit (Na-Dene family – Canada), Cable (2014) proposes that numeral distribution always allows participant-distributive and event-distributive scenarios to be licensed, similar to what is found in Karajá and Karitiana.

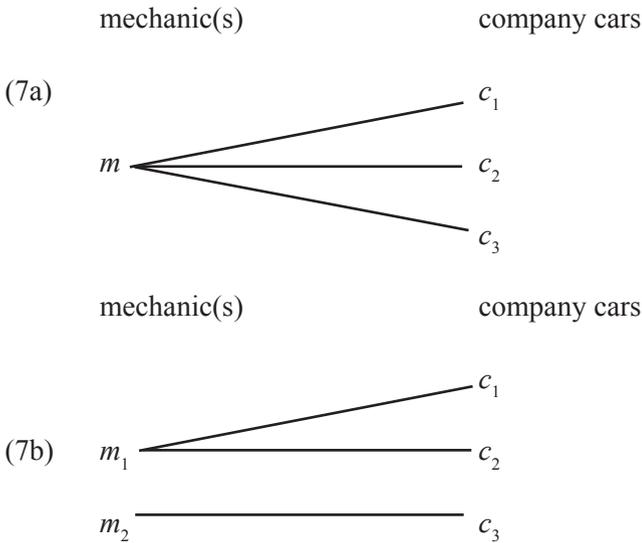
Amongst other facts, I show here that the syntactic attachment site of the distributive numeral has no effect upon the kind of distributive scenario the sentence describes. That is, no matter whether the distributive numeral is adverbial or adnominal, the sentence may be true in either participant-distributive or event-distributive scenarios. (CABLE, 2014, p. 8)

Cable (2014) proposes two semantic denotations, based on a scenario description test, a distributive-adverbial and a distributive-adnominal, for numeral morphology in Tlingit. However, the author argues that both denotations can generate distributive readings about events as well as individuals. Therefore, it is concluded that the ambiguity generated by distributive numerals is recurrent in the languages that have this morphological strategy to express the distributive property in their statements. Based on the data, the ambiguity of the distributive numerals does not seem to be the result of syntactic aspects, but instead the effect of semantic operations which are peculiar to these elements.

Thus, with the intention to ascertain possible differences in processing between possible (and impossible) readings for the reduplicated numeral in Karajá, we conducted an EEG experiment, as presented in Section 4 below. The results can help explain the cognitive mechanisms during numeral distributive processing. The processing of sentences that present quantifying elements is known to be a very complex procedure, since it requires interface between different cognitive modules, such as the syntactic and the quantificational. Therefore, it is part of cognitive research and neuroscience agenda to distinguish the neural mechanisms underlying access to different aspects of semantic memory during sentence comprehension (JIANG *et al.*, 2008).

3. Theoretical assumptions

All natural languages known to date have mechanisms for quantifying sentences. As stated in the introduction, one of the objectives of this paper is to understand the nature and processing of one of the quantifying elements in Karajá, the reduplicated numeral *sohoji-sohoji*, in order to contribute to understanding how quantification works in natural languages and what are the basic principles governing these elements within human languages. The ability to quantify linguistic elements



Syntactically, each interpretation above is represented by a different Logical Form structure that will be generated from the Nominal Phrase raising that is in the verbal complement position. Therefore, for the interpretation of sentences with quantified NPs in the object position, an abstract syntactic structure is necessary to serve as input for the semantic interpretation of sentences with this type of structural component (MAY, 1977, 1985).

Distributive property is an operation that occurs mainly when the nucleus of the Determiner Phrase is occupied by a distributive quantifier and combined with a predicate. The predicate is then understood as applied to each individual member of the quantified set. Therefore, the operation of a distributive quantifier is to map subevents to each available atom of speech. Obviously, this is a very simplistic definition that does not account for a series of distributive mappings that involves scope ambiguity in doubly quantified sentences, for instance. However, it seems to be a statement clear and simple enough to explain the relationship between DPs that contain distributive quantifiers and events to which they are applied, as in “each monkey ate a banana.” In this sentence, “each” takes the individual members of the restrictor set $\{x: x \text{ is monkey}\}$, with the predicate [ate banana] being applied to each individual member of $\{x: x \text{ is monkey}\}$. Furthermore, a distributive quantifier ends up pluralizing the

event of the sentence in which it occurs, since in a distributive structure, at least two subevents of the available set of events are required, i.e., it requires that the R relationship between individuals and events is greater than 1. Then, for sentence (6) above, the meaning can be interpreted as follows: there is an event *e* and for each individual member of {*x*: *x* is monkey} there is a sub-event *e'* which is part of *e* and such that in *e'* *m* “ate banana”.

Distributive quantifiers such as *each* distribute over individual entities. Therefore, when applied to uncountable noun phrases, such as those occupied by mass names of the substance type, they generate an ungrammatical sentence, as illustrated in (8) below, since it would not be possible to individualize its parts, except under coercion or when a unit of measurement is attached to the noun in question. In Karajá, the distributive *sohoji-sohoji* exhibits the same pattern of behavior, already expected for a quantifier of this nature, given the grammaticality of (9a) and the ungrammaticality of (9b).

(8) *He brought each water.

(9a) *Tii hãdoroó sohoji-sohoji-my r-e-wy-reri.*

3P MURICIUM² UM-PosP 3P-VT-TO.CARRY-PRS:CONT

‘He takes murici one by one’ / ‘He takes each murici’

(9b) **Tii bèè sohoji-sohoji-my rewyreri.*

3P WATER UM-UM-PosP 3P-VT-TO.CARRY-PRS:CONT

‘He takes water one by one’ / ‘He takes each water’

Distributive quantifiers also force a distributive reading of the event (subevents) of the sentence in which it is inserted. Therefore, when combined with collective predicates, they generate ungrammatical sentences (DOWTY, 1987). For example, sentence (10) below requires the predicate “be piled up in the corner of the room” be applied to a set of individuals {*x*: *x* books} in (10a) and (10b), and not to individual members of this set, given the ungrammaticality of (10c) and (10d).

² A Brazilian fruit from the *Byrsonima crassifolia* specie.

(10a) The books are piled up in the corner of the room.

(10b) All the books are piled up in the corner of the room.

(10c) *A book is piled up in the corner of the room.

(10d) *Each book is piled up in the corner of the room.

As expected in typically distributive predicates, distributive quantifiers can be applied normally, as in (11) below. Likewise, if a distributive quantifier is not applied to a sentence whose predicate is distributive, a predicate of this nature would still force a distributive reading, even if it does not contain a DP composed of a nucleus quantified by a distributive element in its structure, as seen in (12 and 13) below.

(11) Every girl slept.

(12) All the girls slept.

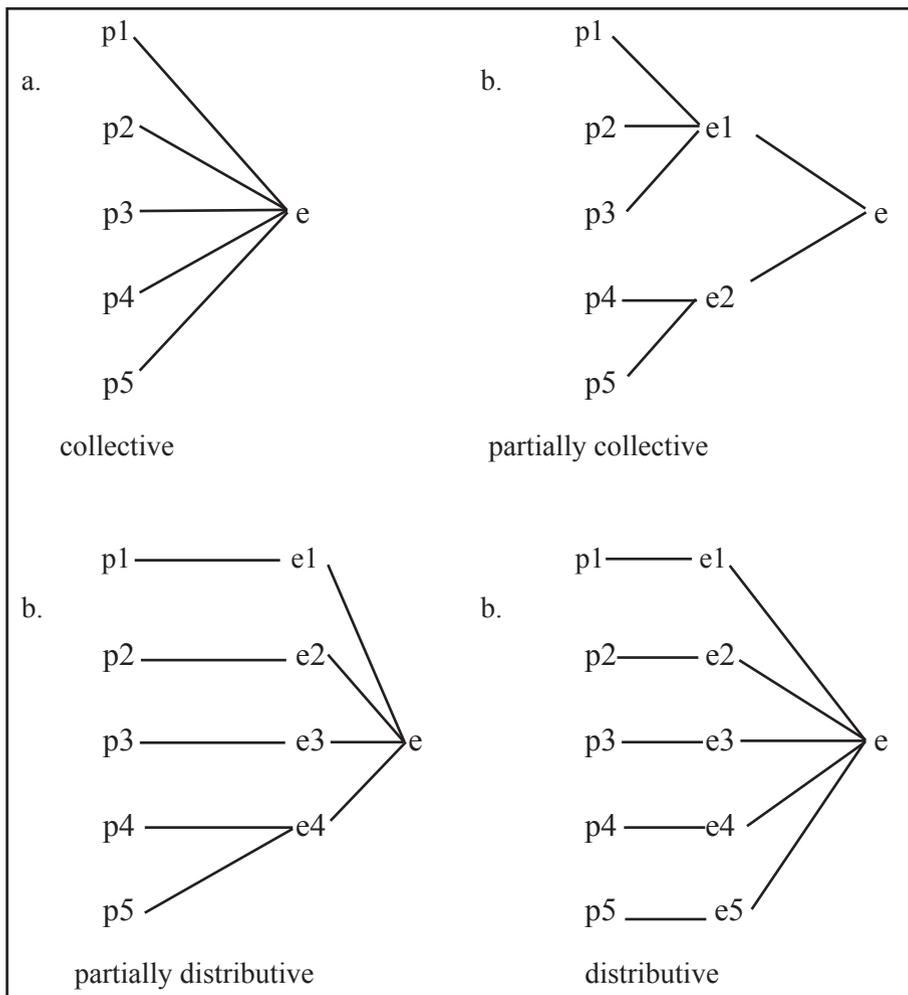
(13) The girls slept.

Thus, according to the above data, it seems that a distributive Reading must operate not only over individuals > 1 , but also over subevents > 1 , thus creating the condition of the differentiation between individuals and subevents – though it is complicated to limit subevents in certain predicates. Non-differentiation of subevents results in a collective reading. However, there are not only collective readings and standard distributive readings. According to Tunstall (1998), distributive-collective event structures form a continuous scale in which partially collective and partially distributive readings are at the middle, while the completely collective and the completely distributive readings are at the ends. Kratzer (2003, 2007) treats intermediate readings of this scale as cumulative. Following Krifka (1998), Kratzer (2003, 2007) introduces the concept of cumulative property as a crucial point for the denotation of predicates in natural languages and argues in favor of the Universal Cumulative property of lexical items that can be conferred in the following section.

“If there is a basic logical-conceptual predicate ‘red’, for example, that is true of my hat and your scarf (two singularities), then the Cumulativity Universal says that that very same predicate is also true of the sum of my hat and your scarf (a plurality).” (KRATZER, 2003, p. 12).

This means that, in simple predicates, a cumulative reading would be available without higher processing costs, as it is inherent and available to lexical items. However, the presence of a distributive operator would result in a preference for a distributive reading. Moreover, the cumulateness inherent in lexical items would also explain the fact that a series of experiments presents evidence of a higher processing cost during the understanding of quantified predicates. This would occur because the presence of quantifying elements would reveal greater complexity for the initial semantic processing of sentences, which would be cumulative by default. The ERP experiment presented in the next section seems to point in this direction. Look at the following diagrams, taken from Tunstall (1998, p. 97), which examine the possible intermediate readings between the standard collective and distributive readings. This scale allowed the author to demonstrate that the main difference between the semantics of *each* and *every* concerns the differentiation condition, which is stronger for *each* (every subevent should be distinct) than for *every* (at least two events must be distinct).

FIGURE 1 – Intermediate readings diagram by Tunstall (1998)



(14) Maria picked up all the folders.

Imagine the following scenario for a sentence like (14) above: there are five document folders on a table such that Maria picked up all these folders. In this scenario, there are a number of possible readings

according to the scale shown above. Reading (a) – **collective** – predicts that Maria picked up all five folders at one time, i.e., there was only one event of “pick up the folders”. Reading (b) – **partially collective** – suggests the folders were picked up in small groups, for example, Maria picked up three folders and then picked up two more folders, thereby generating two subevents of “pick up the folders”, as shown in the above chart. Reading (c) – **partially distributive** – requires there be subevents applied to individual members, but concomitantly allows collective subevents, for example, Maria picked up three folders one at a time and then two folders at one time, i.e., four subevents of “pick up the folders”. Finally, reading (d) – **completely distributive** – requires each folder to be picked up individually, thus generating five “pick up the folders” subevents.

In the following section, we will investigate whether scenarios that have a partial distributive property can be paired with sentences containing the reduplicated numeral *sohoji-sohoji* and what would be the cognitive cost of this procedure. In addition, we aim to demonstrate the acceptability (or otherwise) of different quantified scenarios when the participant was exposed to a sentence operated by *sohoji-sohoji*. The offline and online results corroborate the existence of a gradient of acceptability for different distributive readings, in which completely distributive and singular scenarios are at opposite ends of the scale. We will see that, although the behavioral test results elicited acceptability for the different scenarios, including those that are typically non-distributive, the measurements of time (ms) and amplitude (μ V) seem to provide significant data for our hypothesis that, although different readings may be possible for the *sohoji-sohoji* quantifier, the standard distributive reading is the first one activated by distributive sentences in Karajá whose operator is *sohoji-sohoji*, i.e., the one that requires less cognitive cost during sentence processing.

4. The processing of the reduplicated numeral in Karajá

An experimental design was set up to investigate the processing of the reduplicated numeral, which made it possible to collect online and offline data. The task consisted of pairing different controlled scenarios, which were pictures presented on a computer screen, with distributive sentences recorded in audio. This method is commonly known as picture-matching design.

The objective of this experiment⁴ was to observe whether standard distributive contexts would present processing differences in relation to other quantitative contexts (partially distributive, collective, non-exhaustive, extra-element, and uniqueness) during sentence comprehension with *sohoji-sohoji*. This is because we have observed in previously verified pilot tests that non-prototypically distributive scenarios were accepted by the participants as possible scenarios for sentences containing *sohoji-sohoji*. Therefore, the present experiment aimed to ascertain whether the acceptance of different quantitative scenarios for sentences with *sohoji-sohoji* would be the result of a reflexive process involving independent variables, such as interpretation and accommodation, subsequent to the parsing of sentences, or whether the acceptance of different scenarios allowed by the quantifier *sohoji-sohoji* would be related to the intrinsic semantic properties generated by this reduplicated numeral and, therefore, to be observed at a more reflexive level of sentence comprehension, called online processing.

4.1 Participants

Twenty-two subjects (16 men and 6 women), with corrected or normal vision, normal hearing, right-handed individuals (OLDFIELD, 1971), native speakers of the Karajá language, inhabitants of the Btõiry or Hawalò villages (Ilha do Bananal, TO) participated in this experiment. All participants stated that they had not consumed alcohol or taken any prescription or illicit drug 24 hours before the experiment and had completed elementary education. The mean age of the participants was 26 years.

4.2 Method and Materials

The experiment consisted of presenting the following sequence of events: a picture appeared on a CRT monitor, followed by four audios – corresponding to the one, two, three, and four sequences of a sentence in Karajá. All experimental sentences presented the subject-distributive-locative-verb structure respecting the typological pattern of the language, as in example (15) below. After listening to the audios, the volunteers' task was to decide whether or not the picture would be compatible with the sentence heard. This method is known in experimental psycholinguistics as picture-matching design.

(15) *Krukru sohoji-sohoji òra-tyreki r-y-i-reri.*

PÁSSARO UM-UM TRONCO-POSP 3A-VT-RAIZ-PRS.CONT

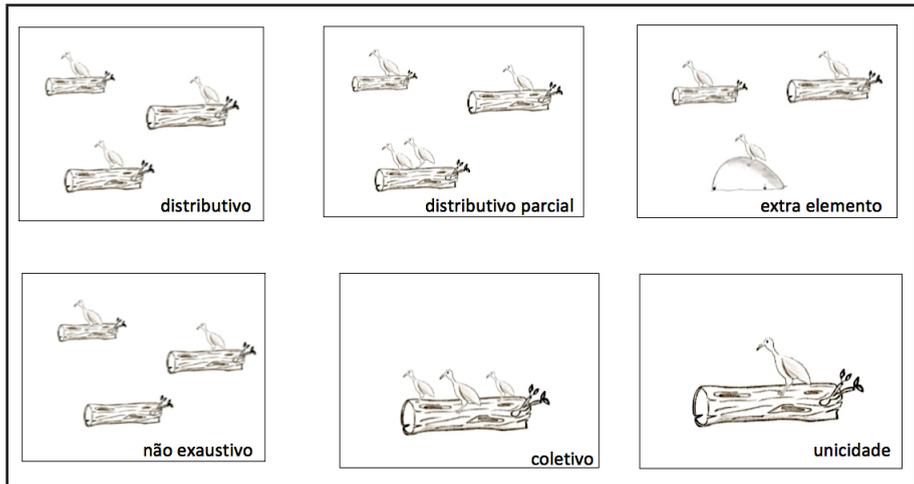
‘Each bird is on the trunk’

Six scenarios were produced for each experimental sentence, namely: (i) **standard distributive**, (ii) **partial distributive**, (iii) **extra-element**, (iv) **non-exhaustive**, (v) **collective**, and (vi) **uniqueness**. The **standard distributive** scenario (i) presented a prototypical distributive property. In other words, the elements of the sets of individuals and the subevents of the set of eventualities were distributed in one-to-one relationships, so that the number of relationships between the sets is greater than 2 and there is a differentiation between the sub-eventualities. The **partial distributive** scenario (ii) showed cardinality failure in the operation of relating individuals and eventualities, since the one-to-one relationship was violated in one of the subevents. The (iii) **extra-element** and (iv) **non-exhaustive** scenarios were complementary. In the first, there was an element of the group of individuals that did not participate in the available subeventuality; in the second, there was a subeventuality that was not related to any individual in question. These scenarios tested the universality/exhaustion property of individuals and eventualities, respectively. The collective scenario (v) was composed of only one eventuality for which all individuals were applied, thus testing whether the property of differentiation between subeventualities was required by *sohoji-sohoji* in the course of processing. Finally, the **uniqueness** scenario (vi) included only one atom of the group of individuals that, in turn, participated in a single eventuality; therefore, it was tested whether a singular context was immediately discarded during the processing of the distributive numeral. For example, in Table (2) below, six scenarios are generated for sentence (15). Pictures were produced with reference to the picture database developed by Karajá designers within the framework of the PRODOCLIN|Karajá project.³ Each picture was drawn with HB pencil on A4 sheet and scanned using 300-bpi scanner equipment. The

³ Karajá Language and Culture Documentation Project, part of the Documentation Program for Indigenous Languages of the Indian Museum (FUNAI), supported by UNESCO.

scenarios were assembled from the scans with the help of Keynote software for Mac OS and exported in JPG format.

FIGURE 2 – Examples of experimental condition scenarios



Sentence: *Kruku sohoji-sohoji orà-tyreki ryireri*. ‘Each bird is on the trunk’

The experiment was programmed in the E-prime software (Psychology Software Tools, Inc. – PST), a commonly used platform for presenting EEG-coupling language stimuli monitored by PyCorder software (Brain Vision LLC). Six versions of the experiment were developed in a Latin square format, so that no participant saw the same experimental sentences, but all the participants saw all the experimental conditions from different sentences.

The lists were presented in both normal and inverse order to avoid list order effects in each version. Starting with the creation of experimental versions distributed among subjects – Latin square design – it is possible to ensure that the same subject never sees the same sentences, yet he/she is exposed to sentences of all conditions. In all, six lists/versions were programmed according to the distribution of the experimental materials in Latin square format. This experimental protocol is important to inhibiting extra-linguistic factors, such as memory, from interfering in sentence comprehension. The distribution of materials among the

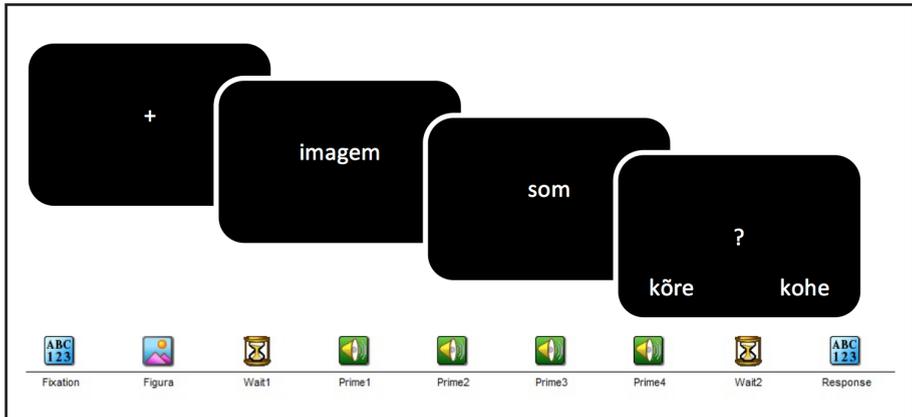
participants also helps the data to be more robust, since analysis between subjects dilutes subjective factors related to the individuals. In addition, the stimuli were pseudo-randomized within each list of each version, with the following restrictions: (i) experimental stimuli were separated by at least one distracting item and (ii) no experimental sentence was followed by another of the same condition.

The total set of materials consisted of 48 experimental sentences plus 48 distractors recorded in audio and 336 digitized and treated pictures. The audios were translated from Portuguese to Karajá with the help of two Karajá language teachers, conferred by a third speaker, and finally recorded by a Karajá consultant. Each experimental sentence was composed of four segments, complying with the Karajá language vocabulary standard, as described in more detail in example (16) below:

(16)	<i>Kruku</i>	<i>sohoji-sohoji</i>	<i>òra-tyreki</i>	<i>r-y-i-reri</i>
	bird	one-one	trunk-on	COP
	subject	distributive	locative	verb
	seg1	seg2	seg3	seg4

The participants were given the task of observing the picture and listening to the sentence that was presented by means of a monitor coupled simultaneously to a PC and an electroencephalography device. At the end of each <image, sound> pair, the task was to assess whether or not the materials were compatible with each other. The participants used a keyboard that had two keys highlighted in green and red. If the participant judged the sentence compatible with the picture, he pressed the green key. If he judged the sentence incompatible with the picture, he pressed the red key.

FIGURE 3 – Screen sequence presented to participants during the EEG experiment



As shown in Figure (3) above, the presentation of the stimulus set began with a cross in the center of the screen for 500ms, followed by the experimental picture/scene for 1500ms. Then the four audio segments were presented. The segments were controlled so that the average amount of time was 1500ms each. The total amount of time between segments was approximately 6000ms. After the sentence was heard, there was a 250ms pause, followed by the response screen, which had a 2000ms time-out, meaning that even if the participant did not press any of the answer buttons, a new data set would be introduced after two seconds. At the end of each experimental sequence, a blank screen lasting 1000 ms was presented before moving to a new <image, sentence> data set.

To determine whether or not there was compatibility between picture and sentence, the volunteers were instructed to push one of the two response buttons, which were balanced between the participants (right/left) to avoid side effects, i.e., to prevent those who already associated the right answer with the right side from benefiting from this association and vice versa. Response times were recorded by *E-prime (Psychology Software Tools, Inc.)* software.

EEG signals were continuously recorded by *PyCorder (Brain Vision LLC)* software from a system consisting of a system of 64 electrodes coupled in a plastic cap (*Brain Products GmbH*) in accordance with the 10-20 extended system (GOMES, 2014). The electrodes were

allocated according to the international standard (*International System Locations*), including five electrodes along the median line (FPz, Fz, Cz, Pz, and Oz) and sixteen lateral/temporal channels, eight for each hemisphere of the brain (FP1/ FP2, F3/ F4, F7/ F8, C3/ C4, T3/ T4, T7/ T8, P3/ P4, and P7/ P8). In addition, another forty-three channels of the extended 10-20 system were used (AF3/ AF4, F1/ F2, F5/ F6, FC1/ FC2, FC3/ FC4, FC5/ FC6, FT7/ FT8, C1/ C2, C5/ C6, CP1/ CP2, CP3/ CP4, CP5/ CP6, TP7/ TP8, P1/ P2, P5/ P6, P7/ P8, PO3/ PO4, PO5/ PO6, PO7/ PO8, CB1/ CB2, and O1/ O2).

The electrodes were referenced online by the ground electrode, located in the frontal region, and later re-referenced by the mean of the mastoid channels (TP9/TP10 – left/right). Impedances were kept below 10 k Ω . The EEG was amplified and scanned at a frequency of 1000Hz. After recording, the data sampling frequency was maintained at 500Hz and the low-pass filters of 30Hz and high-pass of 0.1Hz were used, a method used by the ACESIN laboratory in accordance with international research standards (TANNER *et al.*, 2016). Subsequently, ERP averages were generated for each subject per electrode at times ranging from 200ms to 1200ms in relation to the onset of the target stimulus. Times characterized by blinking or excessive muscle movement were automatically rejected by Vision Analyzer Software (Brain Products GmbH), a program used for filtering and averaging recorded data.

The extraction of ERPs is an electromagnetic cognitive evaluation tool whose application has been extended to linguistic research in the neuroscience of language field since the 1980s, having been developed from the EEG invented by German psychiatrist Hans Berger and used with humans since 1929.⁴ The ERP technique is regarded as very accurate because it registers the central nervous system's responses to motor or sensorial stimulation with millisecond precision and is thus considered an online measurement technique. The EEG provides a continuous temporal sampling of the brain's electrical activity. Such sampling may be

⁴ Like the methods used in experimental psycholinguistics, the ERP technique is also non-invasive, thus allowing it to be used in studies on cognition and language in different human groups, of different age groups, with or without pathologies, and without harm or strain for the participant. To execute the technique, all that is needed is the equipment and software for the gauging, treatment, and statistics of the components that one wishes to investigate.

related to different types of linguistic phenomena and is monitored with the help of specific software by an EEG device attached to a computer that presents the experiment's participants to a certain pre-established linguistic task. After picking up the electrical signals produced in the cortex and monitoring them through the EEG, they must be treated (filters are applied) so that the data can be analyzed (GESUALDI; FRANÇA, 2011). Filtering is an essential, inseparable part of ERP research.

ERPs are composed of a sequence of waves characterized by their latency, amplitude, and polarity. ERP usually has an instantaneous value of 10 to 1000 times less than the background EEG and, therefore, cannot be visualized. To be visualized, the average of several times needs to be performed⁵. This procedure is justified by assuming spontaneous EEG as a zero mean Gaussian white noise and ERPs as the only responses that are actually synchronized with the stimulus. In this way, the effect of averaging is to increase the signal-to-noise ratio (SNR), thus allowing the visualization of the specific effect of the stimulus, in this case, linguistic. (MARQUES, 2011, p. 30)

Thus, averaging is a mathematical treatment of the data used to consolidate the average of the same segment around a given time-coupled moment for all of the stimuli. This treatment aims to exclude random noise that is concurrent with the signal related to the event being studied, such as those resulting from muscular movements, eye blinks, or electrical interference caused by some equipment in the testing room. Since the noise is random, the median after averaging would tend to be zero. Therefore, at the end of this “cleaning” operation, only the signal related to the linguistic investigation will emerge, meaning a large average (or comprehensive average) of the event windows linked to the display time of the stimuli developed for this investigation (LAGE, 2005).

The N400 component was discovered by Kutas and Hillyard in the late 1970s. The researchers developed an experiment to elicit the P3 component (already widely investigated at the time for word recognition) by investigating the role of context in word recognition in sentences such as “eyebrows” in “He shaved off his mustache and

⁵ Times or events are working windows in the temporal continuum of the EEG demarcated for the study.

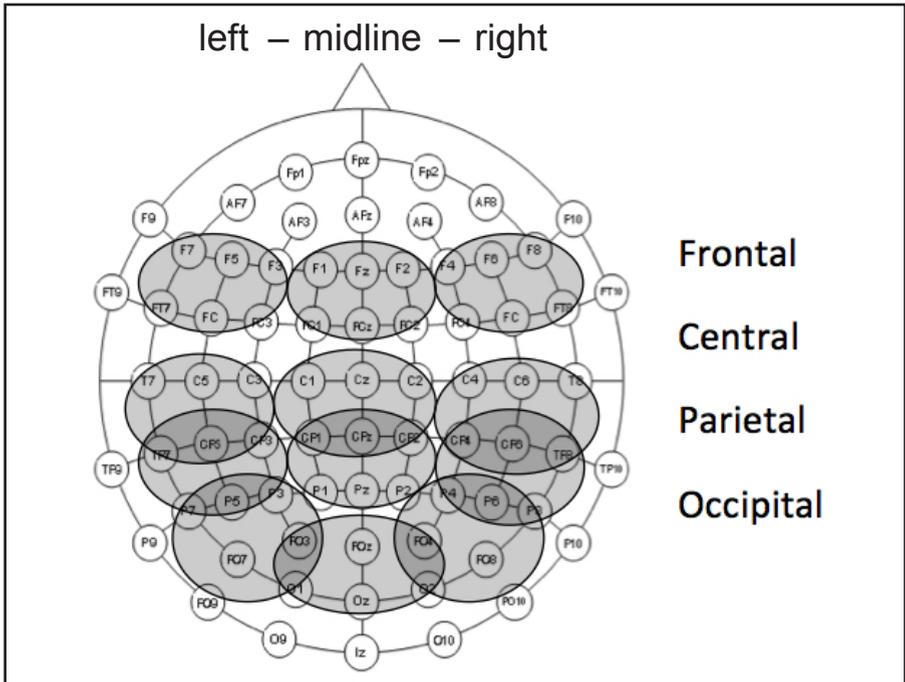
eyebrows.” The results showed a small negative component rather than a positive one as initially expected. To more clearly observe this effect, the authors applied a new experiment with sentences containing a word that, instead of being unanticipated by the context, was semantically anomalous, sentences like “He shaved off his mustache and city” and “I take my coffee with cream and dog”. In the second experiment, the negativity effect was more pronounced, since it elicited very high negative polarity peaks around 400 ms in the course word presentation (KUTAS; HILLYARD, 1980). Since then, this has been the most studied component in the language field, as it is mainly related to the semantic content of the words per se and lexical access theories, since it seems to provide clues on how words are stored (FRANÇA *et al.*, 2008). In Brazil, the study of the N400 component was pioneered in the doctoral thesis of Aniela Improta França (FRANÇA, 2002) entitled *Linguistic concatenations: a study of different cognitive modules in acquisition and in the cortex*. Topographically, N400 tends to exhibit increased activation in the left central and parietal regions of the brain.

Another component studied by language neuroscience, P600, was more recently elicited by Osterhout and Holcomb (1992) and, since then, has been related to syntactic anomalies and incongruities, and can be elicited in sentences of type **The broker persuaded to sell the stock*, in contrast with well-formed sentences like *The broker hoped to sell the stock*. Gouvea *et al.* (1995) also found evidence of high amplitude in the P600 component for sentences that generate temporally incorrect analyses (known in psycholinguistics as garden path sentences). P600 has also been elicited in non-linguistic anomalies in music (PATEL *et al.*, 1998) and mathematics (MARTÍN-LOECHES *et al.*, 2006). In the languages area, in addition to sentence violations, the P600 component also seems to reflect syntax violations, such as violation of number, time, gender, and case agreement, as mentioned by Gouvea *et al.* (1995) and investigated by Gomes (2014), Gomes and França (2015), and Gomes and França (2013). The Late Positive Component (or Late Positive Complex) has already been elicited in studies on the quantification of statements (JIANG *et al.*, 2009; KAAAN *et al.*, 2007), which appear to show that quantified sentences present more complex processing, because the effects of quantificational incongruence are captured at a later time than other language phenomena. Kim and Osterhout (2005) evaluate that

this component can be observed in items whose semantic incoherence could lead to a syntactic reanalysis.

With regard to the analyses presented below, the N400 and P600 components of interest were visually inspected and their peaks were automatically recorded by Vision Analyzer. The peak amplitude and voltage of the components of interest for each of the channels were captured in all segments and conditions. The N400 component was inspected within a 350-550ms window (capturing wide negativity), and the P600 component was inspected within a 550-800ms window (capturing wide positivity) after the onset of the target stimulus. Grand averages were formed by the ERP data of each participant's interest.

FIGURE 4 – Regions of interest based on anatomical proximity



The dependent measures, i.e., the amplitude and latency of the N400 and P600 components, were analyzed by analysis of variance (ANOVA) and parametric tests (T-test). ANOVAs were applied between

the *Condition* (distributive, partial distributive, collective, extra-element, non-exhaustive, and uniqueness), *Laterality* (left hemisphere, right hemisphere, and midline), and *Anteriority* (frontal, central, parietal, and occipital regions) factors. Parametric tests were performed to check any significant difference between the standard distributive condition and other experimental conditions. Therefore, we performed four distinct analyses based on the ERP data:

- i. ANOVA of the peak amplitudes of the waves for the N400 component by condition, laterality, and anteriority. The t test contrasting the means obtained with those of the standard distributive condition.
- ii. ANOVA of the peak amplitudes of the waves for the P600 component by condition, laterality, and anteriority. The t test contrasting the means obtained with those of the standard distributive condition.

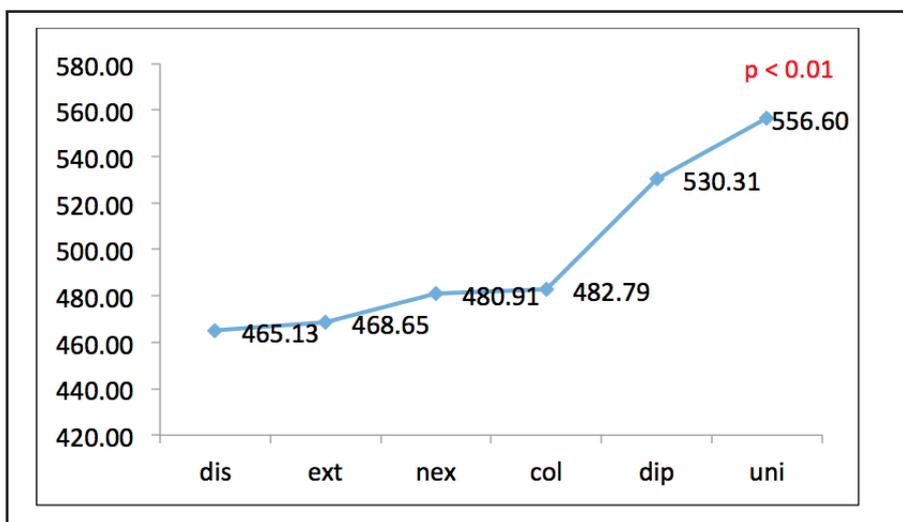
Statistics based on the N400 and P600 components are the most discussed measures in the neuroscience of language literature, and therefore represent the best starting point for an experiment of this nature. Moreover, studies that investigated different aspects of the quantification of statements elicited N400 in quantified sentences, evidencing a higher initial processing cost for statements that involve procedures such as counting and individuation (CHIARELLI *et al.*, 2011). Late positivity effects were also observed, interpreted by Jiang, Tan & Zhou (2009) as an effort to integrate quantifying elements into sentences that present quantificational inadequacy. Therefore, for the analyses presented below, we extracted the amplitude of the N400 and P600 peaks in all experimental conditions from the large means of segments 2 and 3 of the manipulated sentences.

4.3 Behavioral Data

An analysis performed by means of a t test of participants' reaction times (RTs) allowed us to observe which conditions presented the highest average decision time in relation to the standard distributive condition. The results show that only the **uniqueness** condition ($t = -2.585$, $df = 21.182$, $p\text{-value} = 0.0172$) elicited significantly higher

RTs than the distributive condition, which suggests a higher cognitive cost for the decision in this condition. The analysis of raw RTs shows that the uniqueness scenario caused the greatest feeling of strangeness during the pairing task with a sentence containing the distributive numeral. Other conditions did not present significant differences, as can be seen in Graph (1) below.

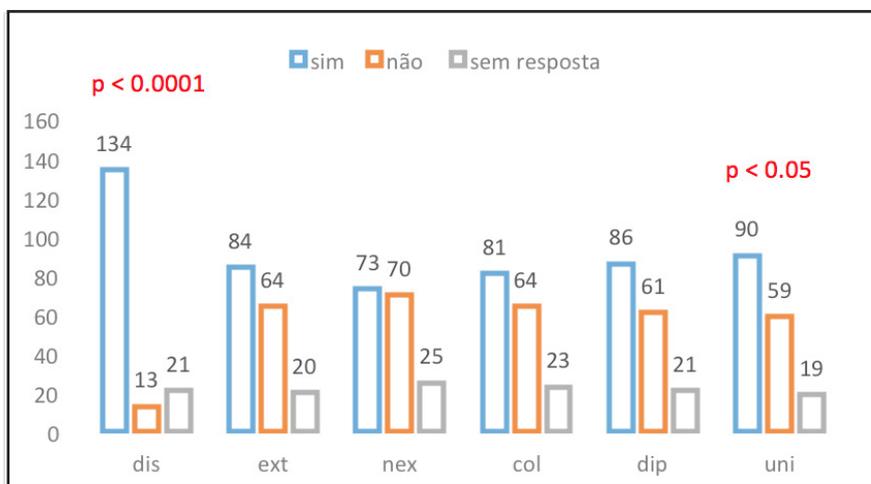
GRAPH 1 – Raw RT per experimental condition



In general, there is a tendency to accept all conditions as compatible scenarios for *sohoji-sohoji* sentences, as shown in Figure 2 below. However, after performing chi-square tests for each condition, it was observed that only the distributive (X-squared = 99.5986, df = NA, p-value = 0.0004998) and uniqueness (X-squared = 6.4497, df = NA, p-value = 0.02249) conditions showed a significant difference between “yes” and “no” answers. The fact that the distributive condition has the highest acceptance rate confirms our expectations since this is the best scenario for the investigated experimental sentences. The acceptance of the uniqueness condition seems to be the result of a cardinal reading attempt of the reduplicated numeral, because, as we will see in the bioelectrical data, this condition was the one that had the highest cognitive cost during processing. Although the other conditions did not

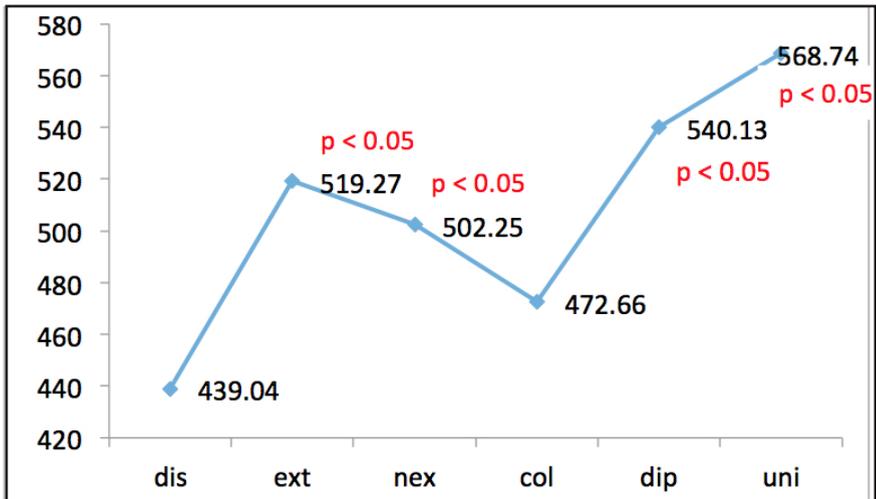
capture a significant effect in this test ($p > 0.05$), we observed that more reflex measures, such as RTs and bioelectrical potentials, reveal there are differences in processing between the different conditions and the standard distributive condition. The 2000ms time-out generated some “no response” items that were excluded from the analyses.

GRAPH 2 – Response rates by condition



Since the rates for “yes” responses were higher in all conditions, the t-test was again applied between conditions to examine only the RTs for the scenarios accepted by the participants, excluding scenarios deemed to be incompatible. **Partial distributive** and **uniqueness** conditions continued to be those with the highest averages, in which case both presented significant differences. However, **non-exhaustive** and **extra element** conditions also attained significantly higher times. The **collective** condition did not present significant RT differences when accepted by the speaker; in this case, the mean was statistically similar to the standard distributive condition. The p-values and averages can be seen in Graph 3 below.

GRAPH 3 – RT of the accepted pairings per experimental condition



These results highlighted the importance of the chronometric data for our analysis, since the acceptability indexes did not prove to be a decisive factor among all the scenarios. Furthermore, although the deviant uniqueness condition has been accepted as a valid scenario for the experimental sentence, the RT analysis shows that this acceptability was performed in a significantly longer period of time than for the standard distributive condition. Without decision time analysis, it would also be impossible to capture the significant distinction between the other experimental conditions, since acceptability remained constant for all scenarios. However, we still cannot get evidence from the RTs that there is a difference in processing between non-exhaustive and extra-element conditions, nor capture if there is a greater cognitive effort to accept the condition uniqueness, as expected.

The behavioral results presented in this section were captured after sentence processing, i.e., offline. We also captured the electrophysiological responses through the ERP extraction technique for online data analysis. The purpose of this technique was to observe the processing of experimental sentences from moment to moment during the temporal

course of sentence comprehension. The results of this method will be presented in the next subsection, point out new reflections regarding the understanding of the processing of the *sohoji-sohoji* distributive, and reveal interesting data on how the acceptability of different scenarios is given.

4.4 ERP Data

The three-way ANOVA (within subjects) model was first used for statistical analysis, which included the experimental condition, laterality, and ROI factors, with correction of the values presented (Greenhouse-Geisser). After the data had been treated, the N400 component showed a significant main effect for the condition factor ($F(5.25) = 7.56$ $p < 0.0000003$) in the second segment of the experimental sentences, corresponding to the *sohoji-sohoji* reduplicated numeral. Conversely, the P600 component elicited a main effect for the condition factor ($F(5.25) = 14.7$ $p < 0.0017565$) only in segment 3, corresponding to the locative postpositional (PP) phrase. No major effects were found for the ROI ($F(11.55) = 1.19$ $p < 0.3472345$) or laterality ($F(2.10) = 1.90$ $p < 0.2238272$) factors in the P600 component nor for the ROI ($F(11.55) = 1.41$ $p < 0.195080$ $p < 0.2820406$) or laterality ($F(2.10) = 2.55$ $p < 0.1680510$) factors in the N400 component. Nor did we find any interaction among the mentioned factors. Therefore, it seems that only the condition factor appears to influence the analyses between the results. Consequently, an analysis was carried out including all ROIs for this article. In the future, other analyses may be performed to confirm the trends presented here.

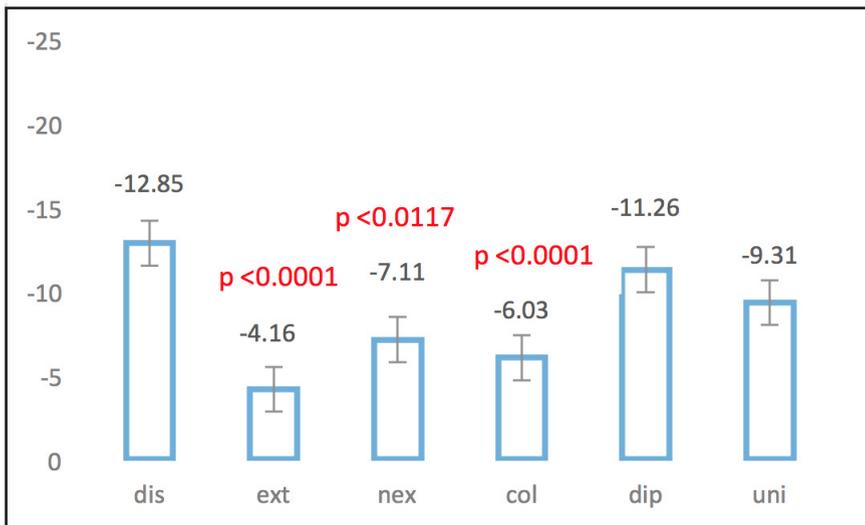
Subsequently, a t test was applied that contrasted the standard distributive condition with the other experimental conditions. Presented below are the results for the amplitude measurements (μV) of the ERP data of the Grand Averages for the N400 and P600 components.

4.4.1 N400 component – negativity after onset of *sohoji-sohoji*

The N400 component, ascertained after the start of the segment containing the *sohoji-sohoji* reduplicated numeral, presented significant results when comparing the standard distributive condition ($-13 \mu V$) with the **non-exhaustive** ($-7 \mu V$, $t(120) = 2.56$ $p < 0.0117$), **collective** ($-6 \mu V$, $t(120) = 4.88$ $p < 0.0001$), and **extra element** ($-4 \mu V$, $t(120) = 6.13$

p < 0.0001) conditions. The uniqueness and partial distributive conditions did not present significant differences.

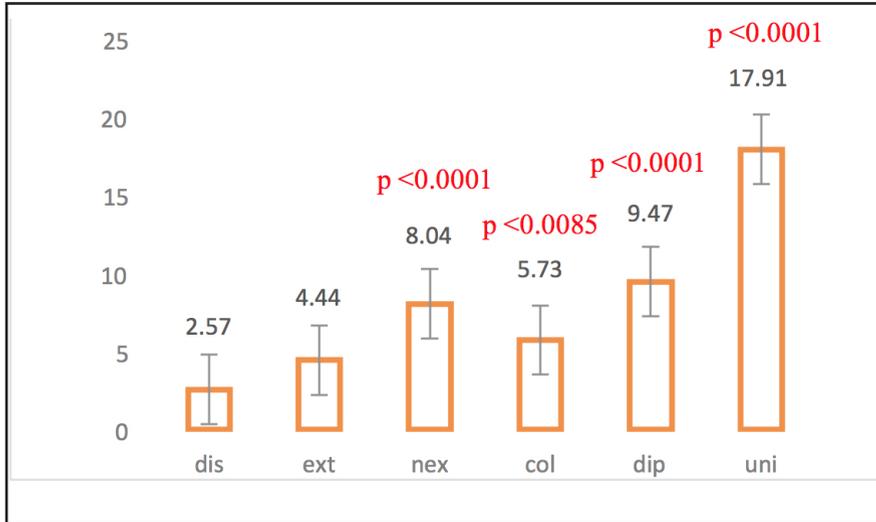
GRAPH 4 – N400 component amplitude averages per condition



4.4.2 P600 component – late positivity after onset of locative PP

The P600 component, ascertained after the start of the segment containing the locative PP of the eventuality, elicited a significant difference between the **uniqueness** (18 µV, t (118) = 17.75 p < 0.0001), **partial distributive** (9 µV, t (118) = 4.91 p < 0.0001), **non-exhaustive** (8 µV, t (118) = 4.17 p < 0.0001), and **collective** (6 µV, t (118) = 2.67 p < 0.0085) conditions when compared to the standard distributive condition (3 µV).

GRAPH 5 – P600 component amplitude averages per condition



4.5 Discussion of results and analyses

The aim of this experiment was to search for a microscopic understanding of the processing of the *sohoji-sohoji* distributive, using previously developed contexts that were provided by scenarios controlled by a specific design, as described at the beginning of this section. The method made it possible to analyze electrophysiological data captured by the EEG during sentence processing, measured online, that would reflect the cognitive cost during sentence processing when paired with different quantitative scenarios. The data were then filtered for extraction and analysis of the amplitudes (μV) captured from the delineation of N400 (250-500ms window) and P600 (550-800ms window). The results indicate that the acceptability of different experimental conditions in behavioral data does not reflect the cognitive cost for pairing between the different scenarios and the reduplicated numeral, since the participants tended to accept all the conditions. However, there was a tendency for slower reaction times (RTs) for scenarios deviating from the distributive

pattern, namely, those that allowed intermediate and non-pluralized distributive readings. Therefore, the results confirm the earlier hypothesis that the processing of cumulative readings, although possible, would be more costly than non-cumulative sentences when there is an element that delineates the quantificational scope, in this case the *sohoji-sohoji* reduplicated numeral. As discussed below, online data suggest that the properties of cumulative property, distributive property, and cardinality exert effects during sentence processing with *sohoji-sohoji* in Karajá. This evidence is in line with current theories on quantification processing and the semantics of distributional numerals, which have been observed in other natural languages.

With regard to ERP amplitude data, N400 elicited lower amplitudes for the **collective** and **extra-element** conditions in segment 2, corresponding to the *sohoji-sohoji* reduplicated numeral. The P600 component elicited higher amplitudes for the **non-exhaustive**, **collective**, **partial distributive**, and **uniqueness** conditions in segment 3, corresponding to the locative PP. These results seem to reflect the cognitive cost of processing the reduplicated numeral against the different scenarios presented during the experiment. The N400 component reveals that the prototypical distributive scenario, that for which there is an individual for each eventuality, has the highest average amplitude, followed by the partial distributive condition. Such evidence indicates that the processing of *sohoji-sohoji* would engage the individuals' cognitive system in a distributive operation requiring a higher processing cost at the reflex level. As for the P600 measurement, registered in the locative segment following *sohoji-sohoji*, the conditions with larger amplitudes are those that break the event-distributive reading, revealing that it is after the start of the pairing procedure of the element in the eventuality that measurements emerge that would show which properties would be more easily processed in distributive numeral sentences.

The **uniqueness** condition was the one that elicited higher mean values in both the P600 (18 μ V) component and the raw behavioral data (557 ms) and filtered by acceptability (568 ms), which suggests the reduplicated numeral requires the pluralization of the statement with respect to the relations established by the individuals and available

eventualities, in agreement with what is expected by distributive elements.

The **partial distributive** condition also presented a high amplitude average for the P600 component (9 μV) and was also the second highest measurement of the reaction times filtered by acceptability (540 ms). These results indicate that intermediate readings, although accepted, present a high processing cost. In this case, as *sohoji-sohoji* is formed from the reduplication of the number one, these results can also be analyzed as coming from the numerical cardinality failure presented by the presented scenario, i.e., a problem in the counting operation.

The **non-exhaustive** condition elicited the P600 component (8 μV) and significant reaction times filtered by acceptability (502 ms). In this condition, an element of eventuality was not in an R relation with any individual in the scenario. The high processing cost shows that the interpretation of sentences with *sohoji-sohoji* requires that the scenario subevents preferentially operate under the scope of a universal quantifier.

Despite eliciting the P600 component (6 μV), the **collective** condition presented lower mean values than did the other conditions mentioned above. Moreover, no effect was captured in the behavioral data. This effect could possibly have been potentiated if events were verified with transitive verbs that allowed a collective and not only a locative reading of the event. This hypothesis is supported by the fact that the participants accepted the collective scenario presented after full sentence processing without significant cognitive costs, since no significant effects on the RTs were found.

The **extra element** condition did not show significant differences in the P600 component, i.e., a scenario in which an individual who did not participate in the event in distribution did not affect the positive amplitude (4 μV). However, in the behavioral results filtered by accuracy (519 ms), this condition presented significant results, indicating that, on a more reflexive, online level, the universality over events is less costly than the universality over individuals during the *sohoji-sohoji* distributive operation. The difference found in the RT may have been the effect of a reinterpretation of post-sentence processing.

Regarding the significant results found in N400 component for the **extra element** ($-4 \mu\text{V}$), **non-exhaustive** ($-7 \mu\text{V}$), and **collective** ($-6 \mu\text{V}$) conditions, one proposal is that, under these conditions, there is a greater delay for engagement of the memory system in computing the distributive property, since lower means reflect less cognitive effort. Therefore, it is appropriate to propose that these results reflect that the semantic incompatibilities of universality and differentiation present in these scenarios would lead the cognitive system to initially not engage in the verification of compatibility of such scenarios, procrastinating verification to a later processing moment, as verified by the results in the P600 component and in the behavioral data. The negativity elicited by the N400 component for the **standard distributive** ($-13 \mu\text{V}$) and **partial distributive** ($-11 \mu\text{V}$) conditions seems to be a result of the incrementality effect of the quantifiers, which would be interpreted in stages, as is the case with negation. (URBACH; KUTAS, 2010; e URBACH; DELONG; KUTAS, 2015). The negativity elicited for the **uniqueness** condition ($-9 \mu\text{V}$) may be a consequence of semantic implausibility in comparing a non-pluralized scenario with a distributive sentence, which would generate a participant's neurophysiological response reflecting the shock in processing a sentence operated by a distributive numeral after being exposed to a singular scenario, which should be immediately discarded.

5. Conclusion

This paper aimed to demonstrate how the availability of ever more sensitive techniques for discovering how the study of language processing can enrich traditional field research and formal language analysis. In this interdisciplinary perspective, the EEG experiment to extract ERPs aimed to investigate how participants would process distributive sentences with *sohoji-sohoji* reduplicated numerals when exposed to scenarios that presented intermediate readings, rather than prototypically distributive ones.

The results demonstrate that the N400 and P600 components were elicited when there were incompatibilities between scenario and sentence. The reanalysis of conditions that violated plurality, cardinality, differentiation, or exhaustiveness was captured by a wide positivity in the P600 component, elicited from elements that occurred as locative

phrases of events. At a previous time, a wide negativity effect on the N400 component was captured under conditions that engaged participants in a quantitative procedure (standard distributive and partial distributive) or immediately dismissed the plausibility between the scenario and the distributive sentence (uniqueness).

These results complement those found in offline chronometric measurements that point out that *sohoji-sohoji* could or would act as an ambiguous quantifier, allowing more than one reading to be licensed for the sentence (event-distributive or participant-distributive readings) or license different scenarios of verification for the same reading, in a perspective close to that proposed by Cable (2014), who proposes a semantic denotation capable of accounting for the different verification scenarios inherent to the distributive quantifiers in Tlingit whose resolution will depend mainly on the context.

In addition, if Choe's (1987) proposal that the distributive operator is a universal quantifier, which has both a constraint set and a scope set, is taken into account, the *sohoji-sohoji* reduplicated numeral as a default would put the eventuality in the set in restriction, i.e., the event-distributive reading would be less costly. This is because the universal operator (\forall) should be applied to the members of the set of eventualities. In other words, the non-exhaustiveness of the elements of the sentence's eventuality would be the reading that would present higher cognitive cost, because it needs reanalysis, as confirmed in the neurophysiological data.

Distributive property is a knowledge that is part of human cognition and found in different natural languages through various morphological and syntactic mechanisms. It is hoped that this work will give visibility to the phenomenon of numeral distributions found in Amerindian languages that may contain relevant properties for the study of areas pertinent to linguistics and, therefore, knowledge of the faculty of human language.

Bibliography

CABLE, S. Distributive numerals and distance distributivity in tlingit (and beyond). *Language*, Washington, Linguistic Society of America, v. 90, n. 4, p. 562-606, 2014.

CHIARELLI, V.; YAGOUBI, R.; MONDINI, S.; BISIACCHI, P.; SEMENZA, C. The syntactic and semantic processing of mass and count nouns: an ERP study. *PLoS ONE*, San Francisco, CA, v. 6, n. 10, p. 1-15, 2011.

DOWTY, D. Collective Predicates, Distributive Predicates, and All. In: ESCOL, 3rd, 1987, Columbus. *Proceedings...* Columbus: Ohio State University, 1987.

FRANÇA, A. I. *Concatenações linguísticas: estudo de diferentes módulos cognitivos na aquisição e no córtex*. 2002. Tese (Doutorado em Linguística) – UFRJ, Rio de Janeiro, 2002.

FRANÇA, A. I.; LEMLE, M.; GESUALDI, A.; CAGY, M.; INFANTOSI, A. F. C. A neurofisiologia do acesso lexical: palavras em português. *Revista Veredas*, v. 2, p. 34-49, UFJF, Juiz de Fora, 2008.

GESUALDI, A. da R.; FRANÇA, A. I. Event-related brain potentials (ERP): an overview. *Revista LinguiStica*, UFRJ, Rio de Janeiro, v. 7, n. 2, p. 24-42, 2011.

GIL, D. *Distributive numerals*. 1982. Dissertation (Ph.D.) – University of California, Los Angeles, 1982.

GOMES, J. N. *Investigating the dissociation between n400 and p600 effects on the syntax semantics interface: an ERP study*. 2014. Tese (Doutorado) – Universidade Federal do Rio de Janeiro, Faculdade de Letras, Rio de Janeiro, 2014.

GOMES, J.; FRANÇA, A. I. A técnica de ERP: investigando a assimetria sujeito-objeto na interface sintaxe-semântica com EEG. *Letras de Hoje* (Online), Pontifícia Universidade Católica do Rio Grande do Sul, v. 4, p. 152-163, 2015.

GOMES, J. N.; FRANÇA, A. I. The ample semantic scope of minute language computations: an ERP study of words in Portuguese. *Revista da ANPOLL*, Anpoll, v. 34, p. 309-346, 2013.

GOUVEA, A. C.; PHILLIPS, C.; KAZANINA, N.; POEPEL, D. The linguistic processes underlying the P600. *Language and Cognitive Processes*, Taylor & Francis Online, v. 25, p. 149-188, 1995.

HEIM, I. *The semantics of definite and indefinite noun phrases*. 1982. Thesis (PhD) – GLSA, University of Massachusetts, Amherst, 1982.

JIANG, X.; TAN, T.; ZHOU, X. Processing the universal quantifier during sentence comprehension: ERP evidence. *Neuropsychologia*, Elsevier, v. 47, p. 1799-1851, 2009.

KAAN, E.; DALLAS, A. C.; BARKLEY, C. M. Processing bare quantifiers in discourse. *Brain Research*, Elsevier, v. 1146, p. 199-209, 2007.

KIM, A.; OSTERHOUT, L. The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, Elsevier, v. 52, p. 205-225 2005.

KRATZER, A. On the plurality of verbs. In: DÖLLING, J.; HEYDE-ZYBATOW, T. (Ed.). *Event structures in linguistic form and interpretation*. Berlin: Mouton de Gruyter, 2007.

KRATZER, A. The event argument and the semantics of verbs. 2003. Manuscrito. Disponível em: <http://semanticsarchive.net>.

KRIFKA, M. The origins of telicity. In: ROTHSTEIN, S. (Ed.). *Events and grammar*. Dordrecht: Kluwer, 1998. p. 197-235.

KUTAS, M.; HILLYARD, S. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, American Association for the Advancement of Science, v. 207, p. 203-205, Jan. 1980.

LAGE, A. C. *Aspectos neurofisiológicos de concatenação e idiomaticidade em português do Brasil: um estudo de potenciais bioelétricos relacionados a eventos linguísticos (ERPs)*. 2005. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

LYONS, J. *Semantics*. London: Cambridge University Press, 1977. 2 v.

MARQUES, F. *Eletrofisiologia da idiomaticidade em indivíduos com Síndrome de Asperger: um estudo de ERPs*. 2011. Dissertação (Mestrado) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2011.

MARTÍN-LOECHES, M.; CASADO, P.; GONZALO, R., DE HERAS, L.; FERNÁNDEZ-FRÍAS, C. Brain potentials to mathematical syntax problems. *Psychophysiology*, Wiley Online Library, v. 43, p. 579-591, 2006.

MAY, R. *The grammar of quantification*. 1977. Dissertation (Ph.D.) – MIT, 1977.

MAY, R. *Logical form: its structure and derivation*. Cambridge, Massachusetts: MIT Press, 1985.

MULLER, A. Distributividade: o caso dos numerais reduplicados em Karitiana. *Cadernos de Estudos Linguísticos*, Campinas, v. 54, n. 2, p. 223-243, 2012.

OLDFIELD, R. C. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, Elsevier, v. 9, p. 97-113, 1971.

OSTERHOUT, L.; HOLCOMB, P. J. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, Elsevier, v. 31, n. 6, p. 785-806, 1992.

PATEL, A. D.; GIBSON, E.; RATNER, J.; BESSON, M.; HOLCOMB, P. J. Processing syntactic relations in language and music: an event-related potential study. *Journal of Cognitive Neuroscience*, MIT Press, v. 10, n. 6, p. 717-733, 1998.

SILVA, C. O. da. *O processamento do numeral distributivo sohoji-sohoji na língua karajá*. 2016. Tese (Doutorado em Linguística) – Faculdade de Letras da Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

TANNER, D.; NORTON, J. J.; MORGAN-SHORT, K.; LUCK, S. J. On high-pass filter artifacts(they're real) and baseline correction (it's a good idea) in ERP/ERMF analysis. *Journal of Neuroscience Methods*, Elsevier, v. 15, n. 266, p. 166-170, Jun. 2016.

TUNSTALL, S. *The interpretation of quantifiers: semantic and*

processing. 1998. Dissertation (PhD) – University of Massachusetts, Amherst, 1998.

URBACH, T. P.; DELONG, K.; KUTAS, M. Quantifiers are incrementally interpreted in context, more than less. *Journal of Memory and Language*, Elsevier, v. 83, p. 79-96, 2015.

URBACH, T. P.; KUTAS, M. Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. *Journal of Memory and Language*, Elsevier, v. 63, p. 158-179, 2010.

Are bilingualism effects on the L1 byproducts of implicit knowledge? Evidence from two experimental tasks

Efeitos do bilinguismo sobre a L1 são produtos de conhecimento implícito? Evidências de duas tarefas experimentais

Ricardo Augusto de Souza

Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais / Brasil

CNPq

ricsouza.ufmg@gmail.com

Cândido Samuel Fonseca de Oliveira

Centro Federal de Educação Tecnológica de Minas Gerais, Belo Horizonte, Minas Gerais / Brasil

coliveira@cefetmg.br

Abstract: Experimental studies in Linguistics rely on data from human participants performing language tasks. Therefore, understanding the constructs that such tasks tap into is fundamental for the interpretation of results yielded by experimental work. In the present study we address issues brought out by a previously published study based on a timed grammaticality judgment tasks that fails to replicate reported evidence of cross-linguistic interaction effects in bilingual processing of argument structure constructions that are not part of the bilinguals' L1 construction repertoire. Although the timed grammaticality judgment task has been argued to be a valid measure of implicit linguistic knowledge, we review recent psychometric studies that challenge this assumption by showing that this task either does not tap into implicit knowledge at all, or does not tap into it as completely as online processing psycholinguistic tasks do. In the present study, we conducted two experiments with the same pool of subjects. One of the experiments employed an online processing task, and the other employed a timed grammaticality judgment task. In

our tasks, sentences in Brazilian Portuguese that emulated the linguistic behavior of the English resultative construction were the target items. We report results that show a mismatch in the observations yielded by the two task types, with only the online processing task revealing apparent L2 effects on performance in the L1. We interpret our results by suggesting that the locus of cross-linguistic interactions in bilingual language processing is mostly related to implicit processes.

Keywords: implicit knowledge; cross-linguistic influences; psycholinguistic tasks; bilingualism; resultative construction.

Resumo: Estudos experimentais em Linguística apoiam-se em dados oriundos de desempenho de participantes em tarefas linguísticas. Portanto, a compreensão dos construtos abordados por tais tarefas é fundamental para a interpretação dos resultados gerados pelo trabalho experimental. Neste estudo, explora-se questões trazidas por um estudo previamente publicado baseado em uma tarefa de julgamento de gramaticalidade temporizada que não replicou evidências anteriormente relatadas acerca de efeitos de interações translinguísticas no processamento bilíngue de construções de estrutura argumental que não fazem parte do repertório construcional da L1 dos bilíngues. Apesar da tarefa de julgamento de gramaticalidade temporizada ter sido defendida como uma medida válida de conhecimento linguístico implícito, resenha-se estudos psicométricos recentes que põem este pressuposto em dúvida, ao mostrar que tal tarefa ou não captura conhecimento implícito, ou não o captura tão completamente quanto o fazem tarefas psicolinguísticas de processamento online. Neste estudo, conduz-se dois experimentos com a mesma amostra de sujeitos. Um dos experimentos empregou uma tarefa de processamento online, e o outro empregou uma tarefa de julgamento de gramaticalidade temporizada. Nessas tarefas, sentenças em português do Brasil que emulavam o comportamento linguístico da construção resultativa do inglês constituíram os itens alvo. Relata-se resultados que mostram a discrepância de observações geradas pelos dois tipos de tarefa, com somente a tarefa de processamento online revelando os aparentes efeitos da L2 sobre o desempenho linguístico da L1. Interpreta-se os resultados como sugestivos de que o local das interações translinguísticas de bilíngues é majoritariamente nos processos implícitos.

Palavras-chave: conhecimento implícito; influências translinguísticas; tarefas psicolinguísticas; bilinguismo; construção resultativa.

Recebido em 5 de janeiro de 2017

Aprovado em 7 de maio de 2017

1 Introduction

The study of human cognitive processes – which includes language and language processing – is challenged by a central problem for those who conduct it as an empirical scientific enterprise. Cognitive scientists often build up hypotheses about processes and architectures that are not liable to direct observation and measurement. Consequently, a crucial part of the job in this field of scientific enquiry is the establishment of reliable and consistent connections between observable facts (e.g.: overt behaviors, response accuracy and type, response latencies, eye movements, deflections in the registration of electrochemical waves produced in the brain, etc.) and the mental states, processes, and traces that the observable facts are hypothetically assumed to instantiate. Thus, advancement in cognitive science is much dependent on the enhancement of the validity of alleged homomorphisms between measures of observed behaviors and measures of the constructs (or latent traces) of which such behaviors are construed as surface manifestations (WILSON, 2005).

Psychometrics is the discipline that targets the development and employment of methods and techniques for construct validity analyses. In experimental work in Linguistics and Psycholinguistics, the critical importance of construct validity is especially highlighted. In this branch of language studies, experiments typically involve the collection and codification of instances of the behavior of human participants who are asked to engage with controlled language tasks. Among such tasks, the grammaticality or acceptability judgments play a significant role, which is only rivaled by tasks designed to capture the real-time processing of strings of linguistic units. The grammaticality judgment task is construed as an offline measure, that is, a measure that captures the output of completed language processing. However, specifically in studies in second language research, it has been proposed that time-ceiling manipulations for responses in grammaticality judgment would yield measures of distinct types of knowledge repositories and distinct levels of cognitive control (ELLIS, R., 2005; BOWLES, 2011).

Accordingly, higher amounts of time for judgment calls would serve as a task tapping into explicit, declarative knowledge about language facts and more controlled processing, both of which would be unavailable should the task be administered in a mode that gave participants but a few seconds for their responses (usually 4 to 6 seconds). In such strictly timed grammaticality judgment tasks, the measure would be of implicit linguistic knowledge.

The interaction between languages in the bilingual mind, and the necessary control upon it that bilinguals must exercise to function in one of their languages, is a central issue in the psycholinguistics of bilingualism (BIALYSTOK *et al.*, 2009). Such cross-linguistic interactions and their control have been hypothesized to be modulated by cognitive and contextual factors in ways that could respond to differences in ultimate attainment in L2 learning. And they have also been hypothesized to correlate with a number of possible cognitive advantages bilinguals would show over the lifespan, such as enhanced metalinguistic ability, and higher accuracy and speed of executive functions (BIALYSTOK *et al.*, 2009). A number of studies show that cross-linguistic interactions are pervasive in bilingualism.

Cross-linguistic interactions have been shown to emerge among high L2 proficiency bilinguals in the form of higher tolerance for argument structure constructions that would be either rejected or processed with much difficulty by monolinguals. Fernández and Souza (2016), as well as Souza (2014) and Fernández, Souza and Carando (2017), have documented observations of this phenomenon employing tasks that deal with language processing in comprehension and production. Such observations have led the authors to hypothesize that the phenomenon is not restricted to a temporary and highly localized processing lapse. Rather, the authors argue that bilinguals exhibit a certain degree of innovation in their overall linguistic representations, that is, in their overall linguistic competence in both L1 and L2. Fernández, Souza and Carando (2017) propose that such bilingual innovations might be one of the psycholinguistic mechanisms behind long-term, gradual contact-induced language change. But in a study that employed a timed grammaticality judgment task, Souza, Soares-Silva and Silva (2016) did not replicate a similar heightened tolerance for L2-like argument structure constructions in the L1 with a pool of equally high L2 proficiency Brazilian-Portuguese English bilinguals. The authors suggest that their findings could indicate

that bilingual cross-linguistic effects are actually evanescent, failing to last long enough to emerge in the timed grammaticality judgment task. This interpretation is ultimately incompatible with the suggestion of bilingual innovations in overall linguistic representations.

However, recent psychometric studies have challenged the assumption that grammaticality judgment tasks and online processing tasks in experimental second language research can be regarded as measures of similar constructs (VAFEE *et al.*, 2016; KIM; NAM, 2016). Specifically, these studies suggest that even timed grammaticality judgments fail to tap implicit linguistic knowledge, which is the nature of representations that is largely assumed to subsidize fluent, automatic language processing. Rather, these psychometric studies suggest that judgment tasks, irrespective of higher or lower temporal ceilings for response, are tapping into explicit linguistic knowledge.

Based on the assumption that bilinguals tend to have enhanced metalinguistic ability when compared to monolinguals since a very early age (BIALYSTOK, 2001; BIALYSTOK *et al.*, 2009), the research question that motivates the present study is whether it could be the case that a timed grammaticality task and an online processing task are actually capturing different phenomena. In order to answer this research question, we had the same pool of participants perform two experimental tasks. Half of our participants were monolingual speakers of Brazilian Portuguese, and half of them were bilingual speakers of Brazilian Portuguese and English with high proficiency in the L2. One of the tasks consisted of a procedure to measure the cost of online sentence processing: the maze task (FOSTER *et al.*, 2009), and the other task was a timed sentence judgment task. Based on the studies we mentioned above, the hypothesis we sought to test was that a bilingualism effect, namely cross-linguistic influences, would be present in the measure that tapped into implicit knowledge (the online processing task), but it would be absent in the timed sentence judgment task.

In the following section, we discuss the relevance of implicit linguistic knowledge and implicit learning for bilingualism studies, and we also review psychometric studies that evaluate the construct validity of the timed grammaticality judgment task as a measure of implicit knowledge. We then move to the description of the linguistic focus of the present study: the English resultative construction. Although the surface form of this construction overlaps with an argument structure pattern

licensed in Brazilian Portuguese, the resultative reading is only linked to the English reconstruction. In this study, we focus on the behavior of bilinguals in the two tasks with respect to sentences that forced the resultative reading into the surface form available in the Portuguese language. After the linguistic analysis of the resultative construction, we present the details of the study design, and we follow to the presentation of our results. We finish this report with a discussion of our interpretation of the results, and with a conclusion in which we point towards future directions.

2 L2 implicit linguistic knowledge and learning, and its measurement

The hypothesis that the human cognitive architecture is supported by two relatively dissociable systems of learning, and therefore two systems for knowledge storage, has been a central debate in cognitive science for many decades now. Hayes and Broadbent's (1988) distinction of a dual learning system is based on the notion of selectivity. According to the authors (HAYES; BROADBENT, 1988, p. 251), one of the subsystems (or as in their terminology, "modes") is "selective, effortful, and reportable", whereas the other "involves the unselective and passive aggregation of information about the co-occurrence of environmental events and features".

Achieving the capacity to function fluently in a second language, especially after a first language is well established, is by all means a daunting cognitive task. Becoming bilingual after primary language acquisition involves adjustments in linguistic representation and processing that cannot be exclusively traced to any given learning and teaching situation, procedure, or strategy. In our opinion, it therefore comes as no surprise that the hypothesis of a dual system for learning and knowledge representation should provide a useful conceptual tool in psycholinguistic attempts to account for high levels of achievement in second language acquisition.

To the best of our knowledge, the first theoretical account of the nature and development of second language ability to rely on the concept of a dual system of representations was Stephen Krashen's model (for example, KRASHEN, 1994). The model, dubbed the *Input Hypothesis*, actually consisted of five interconnected hypotheses, among which the

“acquisition” versus “learning” hypothesis¹ lies at the core. This core hypothesis predicts that there are two distinct and non-commutable processes subsuming second language development. According to the hypothesis, “acquisition” refers to a process that is necessarily subconscious and incidental, and which leads to tacit linguistic knowledge. On the other hand, “learning” is construed as the output of intentional attempts at building knowledge about the organization and functioning of the L2 linguistic system, that is, it involves explicit metalinguistic formulations of some sort. For Krashen (1994), attainment of L2 competence and fluent L2 performance derives exclusively from acquired (i.e: subconscious, tacit, implicit) linguistic representations. Krashen’s model predicts no interface between the two representational subsystems, a point made clear by the doubts cast by the author on the efficacy of overt grammatical instruction and corrective feedback for the development of L2 proficiency (KRASHEN, 1994, p 50-54).

Whereas Krashen’s ideas were largely committed to the hypothesis that the acquisition of L2 competence was driven by innate mechanisms, other developments in cognitive psychology and psycholinguistics have conceptualized implicit learning and implicit representations in language within a framework that accommodates usage-based perspectives on language acquisition and processing.² Like Hayes and Broadbent’s (1988), Reber (1989) and Winter and Reber (1994) define implicit learning as an individual’s capacity to extract regularities from patterned stimuli in the environment, without consciousness of the learning task or reflective action towards it. Both Reber (1989) and Winter

¹ The other four ancillary hypothesis in Krashen’s model are the “natural order of acquisition hypothesis”(which predicts a fixed order of morpheme acquisition in the budding L2 grammar); the “monitor hypothesis” (which predicts that the ability to produce utterances stems exclusively from acquired L2 representations, whereas learned representations account only for post formulation monitoring); the “input hypothesis” (which predicts that comprehension of meaningful messages is the sole mechanism driving L2 acquisition), and the “affective filter hypothesis” (which predicts that only negative affective dispositions can impede comprehensible input to drive L2 acquisition).

² A key distinction between innatist and usage-based perspectives is the point of view held by the latter that language learning is supported by general cognitive capacities, rather than by a language-specific cognitive module. But it should be noted that, as stated by R. Ellis (2005, p. 143), both perspectives agree on viewing the nature of knowledge that supports fluent second language capacity is primarily implicit.

and Reber (1994) rely on experiments on miniature artificial grammar learning as the empirical base to argue for an inductive, automatic and mostly unconscious cognitive architecture that allows for probabilistic generalizations to be made for newly processed input based on previous experience with instances.

In Winter and Reber's (1994) definition there is no specification that implicit learning should be necessarily conceptualized as strictly incidental, that is, implicit learning is not necessarily cost-free in concern to cognitive resources, namely attentional allocation. The role of attention in language learning as a whole, and its role in implicit learning in particular has been a controversial issue among psycholinguistics-oriented second language researchers. A strong version of the notion of implicit learning could be defined as strictly incidental, as discussed above. This would predict that implicit learning might involve learning without attention. But as reviewed by Schmidt (1995, 2001) and by Robinson *et al.* (2014), research findings have not supported such strong version. Although language learning may take place without overt intention and without availability of any conscious recollections of learning effort, the currently available evidence does not support claims that learning might take place without attention to the linguistic input available to the learner. This distinction is framed by Schmidt (1995) as a separation between learning without awareness (a possible operational definition of implicit learning), and learning without attention, a hypothetical mode of learning whose actual existence has not been substantiated by empirical findings.

Schmidt (2001) argues that attention is an umbrella concept for a multicomponent cognitive function that can include subsystems such as alertness, orientation and processing selectivity (subsuming both activation and inhibition of information). In other words, according to Schmidt (2001), even though attention to linguistic patterns may unfold unavailable to introspection and independently of intentionality, it is qualitatively different from preconscious detection, as it requires controlled cognitive processes. If some threshold level of attention to properties and patterns in the L2 linguistic system is paramount for learning, then the emergence of L2 linguistic representations in the bilingual mind is, to quote Schmidt's (2001, p. 29) words, "a side effect of attended processing" of the L2.

Allocation of attentional resources over language processing routines can in turn also be a cognitive operation supported by implicit

knowledge. N. Ellis (2006a) proposes a view of fluent language use as derived from a rational architecture by way of which speakers optimally and implicitly learn the distributional and associative probabilities of their languages, thus achieving processing efficiency that finely converges against the linguistic input. N. Ellis (2006b) further proposes that L2 learning is driven by the same rational learning procedure, despite the usual shortcomings of naturalistic L2 acquisition – when learners often fail to demonstrate acquisition of highly frequent features of the L2 despite intense exposure to the L2 input. The author suggests that the apparent failure to acquire features of an L2 may stem from implicitly learned attention, which results in the blocking of detection of linguistic cues in the weaker language that compete with key cues in the stronger language.

N. Ellis' (2006b) hypothesis is that one's very language learning history builds up implicit representations that guide language users as to what to attend to in the course of overall language processing. Learned attention is fine tuned to one's language experience, and it may in turn constrain one's capacity to promptly attend to, and therefore come to represent features in a new language. This hypothesis has been empirically tested in studies by N. Ellis and Sagarra (2010, 2011), showing both short-term and long-term blocking effects of having learned to attend to adverbial cues in a known language on refocusing attention to verbal morphology cues when dealing with a new language. The authors argue that their findings pose challenges to explanations of L2 variability that evoke maturational decline in language acquisition capacity, interpreting such findings as evidence that the commonly reported failure of L2 learners to achieve native-likeness across the full range of L2-specific features may actually result from the implicit entrenchment of processing routines from the learners' previous language experience. That seems to be an interesting alternative to the hypothesis of a sudden and biologically determined halt in the brain's capacity to acquire new languages after a given age, as it seems to accommodate both the evidence of age effects on ultimate L2 attainment (LONG, 2013) and the growing body of evidence for human neuroplasticity and continued learning capacity despite increased processing demands over the lifespan (RAMSCAR *et al.*, 2014; PAJAKA *et al.*, 2016).

In view of the relevance of the notion of implicit L2 learning and representation, it is not a surprise that the development and validation of measures of implicit knowledge is a major concern for second language

and bilingualism scholars. R. Ellis (2005) reports a study in which a psychometric operationalization of the constructs of implicit and explicit L2 knowledge was suggested. The study gathered data on 17 English grammatical constructions from 111 participants by way of five distinct tasks: an untimed grammaticality judgment task, a timed grammaticality judgment task, an oral imitation task, a oral narrative task, and a measure of participants' capacity to verbalize linguistic rules (thus a metalinguistic capacity measure). The author reports having found two distinct measures across the tasks in his study. His factor analysis³ yielded results showing that the scores in the metalinguistic capacity measure and in the untimed grammaticality judgment task loaded on one factor (which he describes as explicit knowledge), whereas scores in the timed grammaticality judgment task and in the two oral tasks involving production loaded on another factor (which he describes as implicit knowledge). The fundamental difference between the tasks tapping into each knowledge base, according to Ellis (2005), is the processing pressure imposed by the time constraints of the task at hand.

Ellis' (2005) study also showed a difference between the knowledge pool tapped by the grammatical and the ungrammatical sentences, specifically in the untimed grammaticality judgment task. The author suggests that when participants were given sufficient time to rely on explicit linguistic knowledge, accurate rejection of ungrammatical sentences loaded on the explicit knowledge factor.

Other studies support Ellis's (2005) proposal that timed grammaticality judgment tasks at least partially tap into implicit representations. For example, Bowles (2011) replicated Ellis' (2005) results using the same test batteries adapted for Spanish. Furthermore, Bowles (2011) reports an effect of language learning history that is convergent with the assumption that the tasks described by Ellis (2005) lead to processing that mostly relies on different knowledge repositories. In her study, Bowles included two groups of English-Spanish bilinguals, one formed with participants who learned Spanish mostly through

³ Factor analysis comprises of a set of multivariate statistical methods by way of which it is possible to identify the minimal set of variables capable of explaining variability in a data set (for details, see Lowen & Gonulal, 2015). In psychometrics, factor analyses of test results allow researchers to identify the best-fit model of how constructs account for variability in performance.

classroom instruction, and another formed by Spanish heritage speakers. The classroom learner group performed higher in the tasks tapping into explicit knowledge, whereas the heritage speaker group showed the opposite pattern. The overall pattern of Ellis' (2005) findings were also replicated in Godfroid *et al.* (2015), a study in which sentence reading patterns were observed through eye-tracking in combination with timed and untimed grammaticality judgment tasks. In Godfroid *et al.* (2015), comparisons were made between native speakers and non-native speakers of English performing the two types of judgment tasks. The authors report that the type of reading found in the ungrammatical items during untimed task among non-native speakers distinguishes the nature of the processing of such stimuli from the others. The authors interpret such result as evidence of reliance on explicit knowledge, or some similar kind of controlled cognitive process.

Notwithstanding, in recent years there has been a growing controversy as to whether manipulations of time constraint for judgment calls are sufficient to ensure that the grammaticality judgment task can be taken as a valid measure of implicit knowledge. For example, Gutiérrez (2013) reports a factor-analytic study of both timed and untimed judgment tasks in which only the grammaticality or the ungrammaticality of the sentences, not the time constraint, loaded on two distinct factors. The author interprets these two factors as explicit knowledge for ungrammatical sentences in both timed and untimed judgments, and implicit knowledge for grammatical sentences, again irrespective of time constraints.

Kim and Nam (2016) conducted a study to further verify the nature of representations that may be tapped into by distinct task formats, specifically comparing the timed grammaticality judgment task (which relies on receptive processing) and the oral elicited imitation task (which relies on speech production),⁴ notably two tasks that have been identified as tapping into implicit knowledge in previous work. The authors found that the two task types do not load on the same factor, with the production task imposing stricter demands on performance. Kim and Nam (2016) interpret their results as indicating that although the timed grammaticality

⁴ Trials in this experimental task typically involve the experimenter reading the stimulus out loud to the participant, who then follows up with an oral repetition of the stimulus. For a discussion of the properties of the task and its validity as a measure of implicit linguistic knowledge, see Erlam (2009).

judgment task may be taken as at least partially a measure of implicit knowledge, responses to such task might not cover the ultimate level of complexity in the network of implicit representations that subsidize language processing for production. Such network of representations should include form-meaning pairings, pragmatic entailments, collocational restrictions, and probably – in the specific case of bilingualism – cross-linguistic correspondences. Kim and Nam (2016) argue that their findings reveal that even if the timed grammaticality judgment task taps into some implicit knowledge, it may not tap into the same strength of implicit knowledge as that which actually guides real time language processing.

Vafae *et al.* (2016) also investigated the validity of timed and untimed grammaticality judgment tasks as measures of implicit or explicit knowledge vis-à-vis online processing tasks (namely, self-paced reading and a word monitoring task). Through detailed factor analyses, the authors reject the hypothesis that even a timed grammaticality judgment task is a reliable measure of implicit knowledge. Vafae *et al.* (2016) argue that the very nature of online psycholinguistic tasks, which capture language users sensitivity to violations and other linguistic features as comprehension unfolds, are likely to minimize the chances of access to conscious linguistic knowledge. On the other hand, the authors argue that the very nature of a grammaticality task, which quite clearly leads its participants to focus on linguistic forms rather than on comprehension, is likely to invoke explicit knowledge. According to Vafae *et al.* (2016), even though the imposition of strict time constraints may make activation of explicit knowledge harder, it is not possible to rule out that such knowledge base is somehow at stake in judgment calls.

All in all, while it is largely accepted that fluent L1 and L2 users rely on implicitly represented knowledge to obtain efficient performance in language use, it is not yet clear that any form of linguistic judgment task provides researchers with a reliable measure of such implicit representations, at least in concern to bilinguals' L2 knowledge. Therefore, bilingualism studies that employ psycholinguistic online processing measures may be capturing phenomena that could fail to be promptly comparable with observations from experimental designs relying mostly on judgment task performances. We now pass over to the linguistic focus of the present study, which was specifically designed to explore this issue.

3 The Resultative Construction

In the present study we compare BP-English bilinguals and BP monolinguals behavior towards an English argument structure construction, namely the resultative construction (GOLDBERG; JACKENDOFF, 2004; WECHSLER, 2012; OLIVEIRA, 2016). This construction has as its main characteristic the fact that it expresses resultativity. Sentences (1) and (2), for example, express resultativity because in both we have the idea that <the table> reached the property <dry> as a result of the action <wipe> and this result was not entailed by the verb itself. Nevertheless, only sentence (2) is usually considered an instance of the resultative construction because, as opposed to sentence (1), its meaning is not predictable from its components parts (WECHSLER; NOH, 2001).

- (1) Samuel wiped the table until it was dry.
- (2) Samuel wiped the table dry.

There are many syntactic-semantic structures that have been considered instances of the resultative construction. As suggested by Goldberg and Jackendoff (2004), the sentences (3)-(9), can all be considered part of the resultative construction family. In (3) and (4), we have sentences that have a transitive verb that can co-exist with the internal argument irrespective of presence of the resultative predicate. In (3) the result of the action is described by an AP and in (4) by a PP. (5) and (6) also have a transitive verb. However, different from (3) and (4), they cannot co-exist with the internal argument without the resultative predicate due to semantic restrictions. (5) and (6) also have an AP and a PP as a resultative predicate respectively. In (7) and (8) we have resultative sentences that are formed by an intransitive verb. (7) has an AP as its resultative predicate, whereas (8) has a PP. Finally, in (9) we have a resultative sentence whose internal argument is a reflexive pronoun. Even though these sentences vary in regards to some linguistic properties, they all express results that are not entailed by the verb itself and, hence, are considered instances of the resultative construction.

- (3) Samuel watered the flower flat.
- (4) Samuel broke the bathtub into pieces.
- (5) Samuel drank the pub dry.
- (6) Samuel talked us into a stupor.

- (7) The pond froze solid.
- (8) Samuel rolled out of the room.
- (9) Samuel yelled himself hoarse.

In this study, we are going to focus on sentences that have the syntactic-semantic structure of (2) and (3), which are usually classified as the true resultative construction (LEVINSON, 2007). As we can observe in the aforementioned examples, these resultative sentences are formed by an external argument, an atelic transitive verb, an internal argument and a resultative predicate that generates telicity. The difference in telicity between the true resultative construction and its verb is one of the main characteristics of this construction. At first sight, one may think that any AP that can express result can be part of the true resultative construction. However, the resultative predicate has to be an AP that not only expresses the result of the action, but also indicates the limit of the action, which does not have an implicit endpoint. The relation between verbs and adjectives in this construction involves a homomorphic mapping between the temporal structure of the event described by the verb and the scalar property described by the adjective, as discussed by Wechsler (2012). According to the author, maximum endpoint adjectives seem to be the class that best fit such a role in the true resultative construction since they express the limit of a scale (ex: dry = 0% moisture). In (2), by way of illustration, the table is wiped until it goes down all the way on the moisture scale and reaches the limit <dry>.

Native speakers of English seem to be very sensitive to those restrictions in the resultative predicate. Oliveira (2014) conducted an acceptability judgment task with the magnitude estimation paradigm to observe if participants could distinguish true resultative constructions formed by maximum endpoint adjectives, such as (10) and true resultative construction formed by other adjective types (11). The results indicate that the first type of sentence had a mean acceptability of 0.73 in 0-to-1 scale, whereas the second type had a mean acceptability of 0.43. Therefore, there is empirical evidence that the true resultative construction does impose these restrictions to the resultative predicate and that native speakers are sensitive to them.

- (10) One of the classrooms was very dirty, so Desiree swept it clean.
- (11) ??Tara bought a new table, but her crazy brother punched it broken.

Oliveira (2016) points out a set of other restrictions that have been observed in this structure. Verbs and adjectives that come from romance languages, for example, do not usually form resultative sentences. Moreover, past participle adjectives or adjectives with more than two syllables also tend to be unlicensed. Also, the resultative predicate cannot be topicalized and the construction cannot predicate on its external argument. These restrictions seem to be very peculiar to the English true resultative construction, which makes it an interesting topic for studies about second language acquisition and other phenomena related to bilingualism, where the L2 is English.

There have been many proposals of possible resultative constructions in BP, but they all violate some of the basic rules of the resultative construction, as shown by Oliveira (2016). Most of these proposals include sentences with telic verbs or resultative predicates that are not formed by an AP or a PP. Based on that the author contends that the resultative construction is not part of the BP grammar. Thus, the acquisition of this argument structure construction by BP-English bilinguals can be considered the acquisition of a new construction.

In order to study how the acquisition of the resultative construction can influence the L1, it is necessary to ensure that bilinguals indeed acquire the resultative construction. Oliveira and Souza (2012) and Oliveira (2013) show that BP-English bilinguals with high levels of proficiency exhibit acceptability ratings to resultative sentences indicative of successful learning. The latter study also indicates that bilinguals with lower levels of proficiency may not have learned the resultative construction. Therefore, in order to analyze possible effects of the resultative construction acquisition on the L1, we have to investigate bilinguals with high levels of proficiency.

In addition to comparing monolinguals' and bilinguals' behavior towards the resultative construction, we also analyze how these groups behave towards the depictive construction. The depictive construction has the same surface syntactic pattern observed in the resultative construction, namely NP-VP-NP-AP, but the AP is not mapped to a resultative reading. As argued by Pykkänen and McElree (2006), whereas the AP in the resultative sentence in (12) indicates the result and endpoint of the action, the AP in the depictive sentence in (13) indicates the state of the internal argument during the action. More importantly, the depictive construction is licensed not only Portuguese-English bilinguals' L2, but

also in their L1. For this reason the depictive construction was included in the present study as a control to the elicited behaviors concerning the L2-only resultative construction. In fact, the BP syntactic counterparts of both (12) and (13) have a depictive reading and are licensed, as illustrated in (14) and (15) respectively.

(12) Alice combed her hair straight.

(13) Alice combed her hair wet.

(14) *Alice escovou seu cabelo liso.*

‘Alice combed her straight hair’.

(15) *Alice escovou seu cabelo molhado.*

‘Alice combed her hair wet’.

In the next section we provide details of our design for the empirical component of the present study.

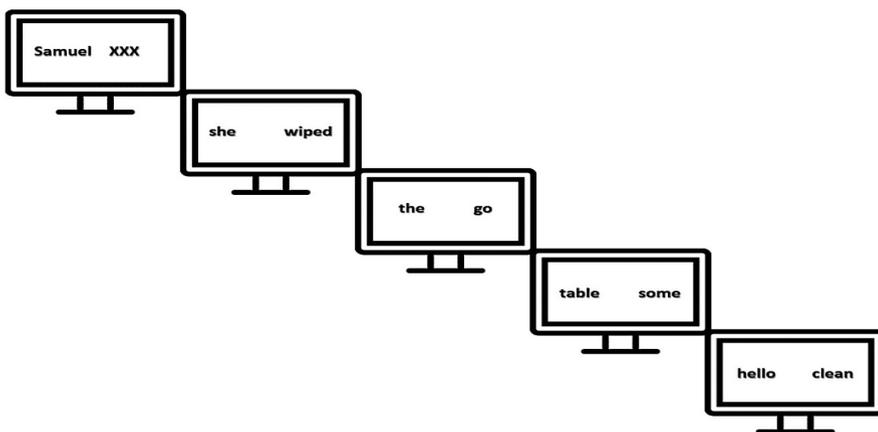
4 Methods

As discussed above, there is growing evidence of a distinction between the constructs tapped into by acceptability judgment tasks and online processing tasks, with only the latter yielding valid measures of implicit processes. The ultimate focus of the present study is to investigate whether such latent construct difference can account for the failure reported by Souza, Soares-Silva and Silva (2016) to replicate bilinguals’ departure from L1 restrictions when processing argument structure constructions in the L1 that are only productive in their L2. It must be recalled that Souza, Soares-Silva and Silva’s (2016) observations were made by way of a timed acceptability judgment task. In order to pursue the present investigation, we planned two experimental tasks (which we refer to as experiment 1 and experiment 2 from now on) to be administered in a within-subjects design. Therefore, we sought to compare the responses elicited by the two types of task from a single pool of participants.

Experiment 1 aimed at measuring the processing cost of sentences that forced an L2-specific construction, namely the resultative construction, into BP by both BP-monolinguals and BP-English bilinguals immersed in the L2. In order to do so, participants performed a maze-task (FORSTER *et al.*, 2009), which is similar to the self-paced reading paradigm in regards to the fragment-by-fragment sentence presentation. The major difference between the two techniques is the fact that in the

maze-task, participants have to choose at each sentence fragment between two options (one leading to a coherent increment to a sentence, the other one fails to do so). Therefore, in the maze-task the reaction times (RTs) of each sentence fragment reflects how long participants take to select the correct option. The main advantage of this method is the fact that it does not require comprehension questions, it does not exhibit spillover effects and it forces an incremental processing. In FIG. 1, we have an example of how the sentence “Samuel wiped the table clean” would be displayed in a maze-task. Each screen the participants see when reading this sentence is represented.

FIGURE 1 – Example of how the sentence “Samuel wiped the table clean” could be displayed in a maze task
Source: Oliveira (2016)



In order to observe possible differences between bilinguals and monolinguals, we compared their RTs for the APs in the target sentences, which are ungrammatical in BP, but not in English. Furthermore, we compared bilinguals’ and monolinguals’ RTs for the APs in control sentences with the depictive construction, which is licensed in both BP and English. Therefore, in this task the dependent variable was the RTs towards the APs, of target and control sentences separately, and the independent variable was the participants’ linguistic profile.

Experiment 2 aimed at analyzing how participants perceived the acceptability of the same constructions from Experiment 1 under time pressure. In other words, participants performed a speeded acceptability judgment task with a 4-second time ceiling. The 4-second time ceiling (i.e.: 4000ms) time-ceiling was based on Souza *et al.* (2015), which reports an exploratory study showing that the 4000ms time-window was about 500ms above the lowest threshold observed for adult and post-secondary education native speakers to make accurate acceptability judgments in their L1. In this task, participants read entire sentences and evaluate how acceptable each of them sound. In order to observe possible differences between bilinguals and monolinguals, we compared the acceptability ratings they assigned to the target sentences, instances of the resultative construction, and the control sentences (as previously stated, instances of the depictive construction). Thus, in this task the dependent variable was the acceptability ratings for the target and control sentences and the independent variables were the participants' linguistic profile and the constructions instantiated by target and control sentences.

4.1 Participants

43 people participated in both Experiment 1 and Experiment 2. Their mean age was 26 and they were college students or had higher levels of education. As tested by Oliveira (2016), performing the speeded acceptability judgment task after a maze-task with similar target structures does not seem to bias participants' behavior as a result of order effects. 27 participants were monolinguals or had only basic knowledge of an L2 and they were all residents of the Belo Horizonte metropolitan area. 16 participants were BP-English bilinguals with high levels of L2 proficiency and were residents of the Boston metropolitan area. These bilinguals had been living in the United States for longer than 10 years, but still considered BP their dominant language. All bilinguals reached the highest level of the Vocabulary Levels Test (NATION, 1990). The Vocabulary Levels Test has been empirically shown to obtain scores that reliably correlate to proficiency test scores based on tasks tapping into grammatical knowledge and comprehension skills in the L2 with Brazilian Portuguese L1-English L2 speakers (SOUZA; SOARES-SILVA, 2015). We therefore understand this test to be an effective diagnosis of bilinguals' levels of proficiency.

4.2 Materials

Both experiments used stimuli in Brazilian Portuguese, the participants' L1. The maze task had 58 experimental items, 10 of which formed the training session. All the sentences were grammatical, except for the 8 target items, whose resultative predicate (AP) is unlicensed in BP. These target items, exemplified by (16), forced the resultative construction using a structure suggested by Oliveira (2013). All items had two clauses. The first one had an NP in the subject position, a verb in the past tense, and a direct object NP with a definite article. The second one had a coordinating conjunction, another verb in the past tense, a clitic pronoun as direct object referring to the direct object of the previous clause, and an ungrammatical AP that forced a resultative reading. The two NPs in each sentence differed from each other in terms of gender in order to decrease the possibility of ambiguous readings. The 8 control items, exemplified by (17), were instances of the depictive construction and their structure were similar to the target items. The only difference between the target and the control items was the interpretation of the AP. More specifically, whereas the AP in the target items had a resultative reading, which made it ungrammatical in BP, the AP in the control items had a depictive reading, which is licensed in BP.

(16) **A criança pintou o papel e o soprou seco.*
 DET kid paint.PST DET paper and it.ACC blow.PST dry
 'The kid painted the paper and blew it dry.'

(17) *A professora preparou o chá e o bebeu quente.*
 DET teacher prepare.PST DET tea and it.ACC drink.PST hot
 'The teacher prepared the tea and drank it hot.'

The speeded acceptability judgment task had 111⁵ experimental items, 15 of which formed the training session. The experimental corpus was balanced in terms of grammaticality so that 50% of the sentences were grammatical and the other 50% were ungrammatical. Thus, it was possible to mitigate possible effects related to the repetition of sentences

⁵ More items were included in the acceptability judgment task than in the maze task because of the different nature of those tasks. For the acceptability judgment task, we believe a wider array of different sentence types are needed in order to keep the target sentences from becoming too salient, and therefore raising participant's awareness towards them.

with similar grammatical status. The 8 target (18) and 8 control (19) items were similar in structure to the target and control items of the maze task as illustrated below:

(18) *O menino pintou a unha e a soprou seca.*
 DET boy paint.PST DET nail and it.ACC blow.PST dry
 ‘The boy painted the nail and blew it dry.’

(19) *A moça preparou o café e o bebeu quente.*
 DET girl prepare.PST DET coffee and it.ACC drink.PST hot
 ‘The girl prepared the coffee and drank it hot.’

In both the maze-task and the speeded acceptability judgment task the items were pseudo-randomized. Such a procedure aimed at avoiding that target and control sentences were displayed in sequence. Therefore, it was also possible to mitigate possible effects originated from the repetition of the same construction and/or the order of presentation.

4.3 Procedures

Participants performed the maze-task (Experiment 1) and, after an interval in which they provided their personal information, they performed the speeded acceptability judgment task (Experiment 2). Both tasks were performed in the same laptop computer. The DMDX software (FORSTER; FORSTER, 2003) was used for the stimuli presentation and the randomization management. In the maze-task, the software recorded reaction times (RTs) for each segment and, in the speeded acceptability judgment task, it recorded the acceptability ratings given to each sentence.

In both tasks participants were introduced to a set of instructions. In the maze-task, the instructions informed the participants that they should form sentences by choosing, from each pair of words, the option that best suited the sentence being formed. In order to select words, participants used the left-shift and right-shift keys, which were highlighted with colored stickers. In the speeded acceptability judgment task, participants were instructed to assess the acceptability of each sentence with a 5-point Likert scale by using the number keys from 1 to 5. This scale has been argued to be the most suitable for this type of task (SOUZA; OLIVEIRA, 2014). Also, they were instructed to judge the sentences based on the order used and the words selected, trying to ignore pragmatic aspects.

Both Experiment 1 and Experiment 2 were fully conducted in BP, in order not to encourage the activation of participants' L2, i.e., to keep them in a monolingual mode (GROSJEAN, 2013). Both tasks were preceded by a training session and participants could ask questions to the experimenter. In order to avoid possible fatigue effects, participants took a break halfway through each task. In the maze-task, participants had 4000ms to read each pair of words and select the correct option. In the speeded acceptability judgment task participants had 4000ms to read and assign an acceptability rating to each sentence. As stated above, the 4000ms time-ceiling was based on the results reported in Souza *et al.* (2015), concerning the time-window within which adults with post-secondary education can make accurate acceptability judgments in their L1.

5 Results

5.1 Experiment 1

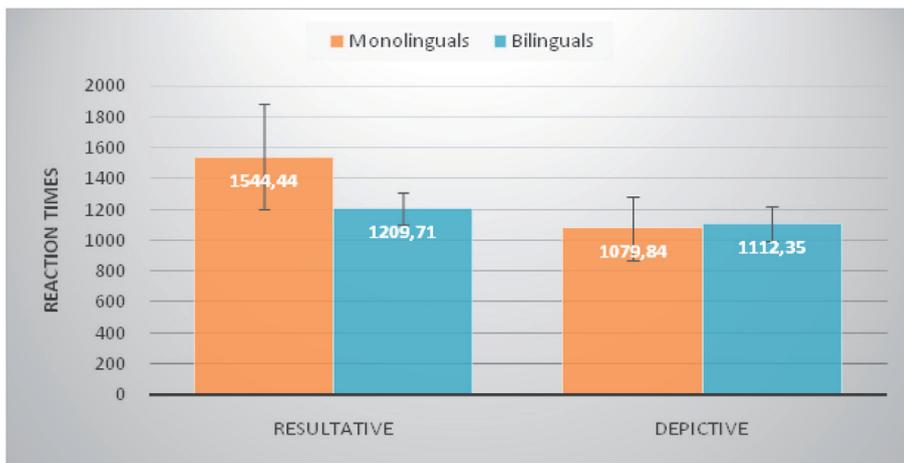
Our hypothesis for Experiment 1 was that bilinguals would exhibit shorter RTs for the AP in the resultative construction in comparison to monolinguals. The rationale is that bilinguals would co-activate both the L1 and the L2 and, in turn, they would be able to process the resultative predicate more easily than monolinguals. Since the control items are available both in the L1 and in the L2, we did not expect that they would yield significant differences between bilinguals and monolinguals.

We tested the maze-task target and control item RTs for normality with the Shapiro-Wilk test. The monolinguals' means for the resultative construction by subjects ($W=.893$, $p=.252$) and by items ($W=.920$, $p=.433$) did not differ from the normal distribution, and neither did their means for the depictive construction by subjects ($W=.948$, $p=.692$) or by items ($W=.959$, $p=.796$). Similarly, the bilinguals' means for the resultative construction by subjects ($W=.874$, $p=.166$) and by items ($W=.908$, $p=.388$) did not differ from the normal distribution, and neither did their means for the depictive construction by subjects ($W=.956$, $p=.770$) or by items ($W=.933$, $p=.544$).

Due to the normality of all the distributions observed, we used the Student's T-test for independent samples to compare bilinguals' and monolinguals' RT means by subject and by items to both target and control sentences. The groups' RTs for the AP in the resultative sentences yielded a significant difference by subjects ($t(39)=3.725$, $p<.001$) and by items

($t(14)=4.732$, $p<.001$), but their RTs for the AP in the depictive sentences did not either in neither the analysis by subjects ($t(39)=-.584$, $p<.563$) nor in the analysis by items ($t(14)=.442$, $p<.665$), as we suspected. The participants' RTs means for both sentence types are illustrated in GRAPH 1. Thus, bilinguals and monolinguals did not exhibit a significant difference as regards the RTs to the APs in the depictive sentences, which are licensed in BP, but bilinguals were significantly faster in relation to the APs in the resultative sentences, which are illicit in BP, but licit in the bilinguals' L2. Therefore, the results suggest that bilinguals exhibit facilitation possibly originated from an access to the L2 representation during online processing of the argument structure pattern not licensed in their L1.

GRAPH 1 – Monolinguals' and bilinguals' mean RTs for the APs in the resultative and in the depictive constructions



5.2 Experiment 2

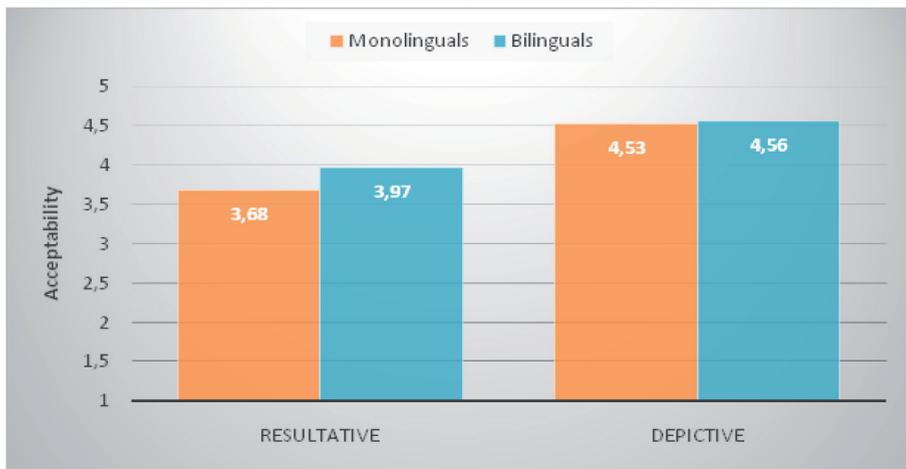
Our hypothesis for Experiment 2 was that bilinguals would not exhibit difference acceptability ratings for the resultative construction in comparison to monolinguals. The rationale is that the timed acceptability judgment task taps into a knowledge different from that involved in the maze-task. As discussed above, our hypothesis is that whereas the maze-task taps into implicit knowledge, the speeded acceptability judgment task taps into explicit knowledge. Since the L2 influence on L1 is an evanescent, temporary and implicit effect on L1 (SOUZA *et al.*, 2016), we do not expect to find this effect on the speeded acceptability judgment

task. Therefore, we do not expect to find differences between bilinguals and monolinguals as for the manner they perceive the acceptability of both the resultative and the depictive sentences.

We tested the acceptability ratings for normality with the Shapiro-Wilk test. The monolinguals' means for the resultative construction ($W=.841$, $p<.001$) and for the depictive construction ($W=.603$, $p<.001$) differed from the normal distribution. Also, bilinguals' means for the resultative construction ($W=.740$, $p<.001$) and for the depictive construction ($W=.415$, $p<.001$) differed from the normal distribution.

We ran the Mann-Whitney test to compare bilinguals' and monolinguals' acceptability ratings. The acceptability ratings given to the resultative construction did not yield a significant difference between bilinguals and monolinguals ($U=9751.5$, $W=16537.5$, $Z=-.262$, $p<.793$). Similar results were observed in regards to the acceptability ratings given to the depictive construction ($U=9930.5$, $W=25861.5$, $Z=-.914$, $p<.360$). The participants' acceptability rating means for both sentence types are illustrated in GRAPH 2. Differently from what was observed in the maze-task, bilinguals and monolinguals behaved similarly, as we suspected. More specifically, bilinguals' behavior did not suggest an L2-to-L1 influence to evaluate the acceptability of the resultative sentences. Thus, the results suggest that the L2 influence on the L1 processing does not last long enough to play a role in participants' metalinguistic analysis regarding the acceptability of an argument structure construction that is L2-specific.

GRAPH 2 – Monolinguals' and bilinguals' mean acceptability for the the resultative and the depictive constructions



6 Discussion

Although the present study was not designed for factor-analytic data treatment, our results do show that a task that taps into more automatic and online processing (the maze task) yields distinct behavior from what was measured in a timed grammaticality judgment task. Therefore, we understand the present study to converge with both Vafae *et al.* (2016) and Kim and Nam (2016) in respect to the fact that psycholinguistic measures are likely to capture features of processing that are dissociable from what is measured by grammaticality judgments, even when the time window for responses in the latter are manipulated so as to impose restrictions on reflective analysis of stimuli by respondents. It must be emphasized once again that the present study explored the behavior of the same participants for each of the two experiments, therefore controlling for the possibility that the variability in the performance of the two task types resulted from individual differences clustered in one of the participant pools.

Following the arguments in Vafae *et al.* (2016), we interpret our observations as reflecting the differences in outcome of processing in a task that involves automatized and largely implicit parsing routines (what we detected through the maze task) as compared with processing in a task that allowed our participants to rely on explicit knowledge of some sort. Such explicit knowledge should not necessarily take the shape of metalinguistic descriptive rules about the linguistic behavior of the resultative construction. Even the awareness of the contrastive readings entailed by the surface forms of sentences like as *Samuel wiped the table dry* and *Samuel limpou a mesa seca* might have been available for inspection for our participants over the timed grammaticality judgment task they performed.

We believe that our sample of Brazilian Portuguese-English bilinguals was composed of L2 speakers who achieved a reasonable degree of automaticity and proceduralization in parsing L2 constructions. First, the bilingual participants in our sample were all classified in the highest possible level in the Vocabulary Levels Test, a score that has been shown to positively and significantly correlate with independent measures of L2 proficiency that include morphosyntax (SOUZA; SOARES-SILVA, 2015; SOARES-SILVA, 2016). Second, those were bilinguals immersed in the L2 sociolinguistic environment for several years. This fact taken together with their explicitly measured proficiency in the L2 makes it highly probable that the bilingual population we sampled is characterized by intense

exposure and productive use of the L2. The performance of our bilinguals in the maze task revealed a processing pattern that diverged from what we observed among monolinguals in the critical items of our experiment, and effects of intense usage of the L2 on the altering of bilingual processing of L1 patterns in relation to monolinguals have been documented elsewhere (e.g.: FERNÁNDEZ, 2003; DUSSIAS; SAGARRA, 2007). Ultimately, as discussed above, there is evidence that it is precisely the abstraction of distributional probabilities in the input that characterizes much of the implicit knowledge that takes the shape of grammar.

As discussed above, Souza, Soares-Silva and Silva (2016) suggested that the non-replication in a timed judgment task of the bilingual cross-linguistic effects found in online processing tasks could be a temporal decay of such effects. However, the considerations we bring out in the present study lead us to suggest a different perspective on this issue. We propose the alternative hypothesis that instead of the consequence of a time factor, the mismatch reported in the authors' study could be a byproduct of the fact that bilingual cross-linguistic interactions may take place mostly at the level of implicit representations and processes, but are actually inhibited and controlled if the task contingencies allow for integration of explicit representations.

Thus, we interpret our observations in the present study as suggesting that the extent to which bilingual cross-linguistic interactions occur is modulated by task type, and that such interactions can be clearer when performance relies on implicit representations, rather than on more controlled processes. Therefore, we assume that the hypothesis that departure from L1 norms may reflect changes in linguistic competence, put forward in Fernández and Souza (2016) and Fernández, Souza and Carando (2017), could be re-stated as changes in the output of procedures that rely on implicit representations. This assumption is subsumed by the notion that linguistic competence does not refer to a stable knowledge repository, but rather to a malleable set of representations that optimally serve the pressures for the efficient resolution of language processing demands at the point of need, and as imposed by the specific contingencies of specific linguistic tasks.

7 Conclusion

Our observations in the present study are convergent with the results reported in Fernández, Souza and Carando (2017) and Souza (2014) in concern to a facilitative bilingualism effect on the L1 online processing of an argument structure construction that is alien in the bilinguals' L1, but productive in their L2. On the other hand, Souza, Soares-Silva and Silva's (2016) observations that such bilingualism effects are not detectable in timed acceptability judgments were also replicated. We interpret these findings in light of recent research findings that challenge the assumption that the reduction of the time ceiling for judgment calls makes the grammaticality judgment task a reliable measure of mostly implicit linguistic knowledge. In other words, we understand our present findings as suggesting that bilingual cross-linguistic effects on the L1 that have been reported in literature might be specifically salient in processes that rely on implicit linguistic knowledge.

This interpretation seems to be consistent with the notion that implicit processes are automatic, whereas explicit processes are likely to be more controlled. Ultimately, the picture that we tentatively put forward here is that inhibition of properties of a language irrelevant for a given processing task will be more effective in the performance of tasks that are less dependent on the immediate outcome of automatic processes, therefore allowing for more language control. This seems to be the case of the acceptability judgment task, even under severe time restrictions. We believe that such tentative picture also seems to accommodate the hypothesis that continued automatic activation of non-relevant language features during online processing might in the long run alter linguistic representations that may be accessed even by explicit processes, thus leading to bilingual innovations that may ultimately drive language change, as proposed by Fernández, Souza and Carando (2017).

A limitation of the present study is the fact that we did not employ a specifically factor-analytic design for the treatment of our data. We based our assumption that explicit knowledge was tapped into by the timed grammaticality judgment task on other research findings that show that bilinguals in general have enhanced metalinguistic capacity. However, we understand that more detailed psychometric validation studies of the constructs tapped into by psycholinguistic tasks are an invaluable future direction for both the psycholinguistics of bilingualism in particular and

for psycholinguistic research at large. Another limitation of this study is the fact that we did not employ any validated measurement of language use and/or dominance profile with our bilingual participant pool. The lack of such an instrument impedes us from holding any conclusive position concerning whether it was indeed the high level of attained L2 proficiency – rather than emerging drifts towards L2 dominance resulting from immersion in the L2 sociolinguistic environment – that better explains the bilingualism effects on L1 processing we did observe in our online processing task. These limitations notwithstanding, we hope the present study is above all an example of the fruitful collaboration that can be established between experimental psycholinguistics and solid psychometric considerations.

Acknowledgements

The authors would like to express their gratitude for grant APQ-02932-14, from the Fundação de Amparo à Pesquisa do Estado de Minas Gerais, which supported the present study.

References

- BIALYSTOK, E. *Bilingualism and development: language, literacy and cognition*. Cambridge; New York: Cambridge University Press, 2001. <https://doi.org/10.1017/CBO9780511605963>
- BIALYSTOK, E.; CRAIK, F.I.M.; GREEN, D. W.; GOLLAN, T. H. Bilingual minds. *Psychological Science in the Public Interest*, Sage Journals, v. 10, n. 3, p. 89-129, 2009.
- BOWLES, M. Measuring implicit and explicit linguistic knowledge – What can heritage language learners contribute? *Studies in Second Language Acquisition*, Cambridge University Press, v. 33, n. 2, p. 247-271, 2011.
- DUSSIAS, P.; SAGARRA, N. The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 10, n. 1, p. 101-116, 2007.
- ELLIS, N. Language acquisition as rational contingency learning. *Applied Linguistics*, Oxford Academic, v. 27 n. 1, p. 1-24, 2006a.

ELLIS, N. Selective attention and transfer phenomena in L2 acquisition: contingency, cue competition, salience, interference, overshadowing, blocking and perceptual learning. *Applied Linguistics*, Oxford Academic, v. 27, n. 2, p. 164-194, 2006b.

ELLIS, N.; SAGARRA, N. Learned attention in adult language acquisition – A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition*, Cambridge University Press, v. 33, n. 4, p. 589-624, 2011.

ELLIS, N.; SAGARRA, N. The bounds of adult language acquisition – Blocking and learned attention. *Studies in Second Language Acquisition*, Cambridge University Press, v. 32, n. 4, p. 553-580, 2010.

ELLIS, R. Measuring implicit and explicit knowledge of a second language – A psychometric study. *Studies in Second Language Acquisition*, Cambridge University Press, v. 27, n. 2, p. 141-172, 2005.

ERLAM, R. The elicited oral imitation test as a measure of implicit knowledge. In: ELLIS, Rod; LOWEN, Shawn; ELDER, Catherine; ERLAM, Rosemary; PHILIP, Jenefer; REINDERS, Hayo (Ed.). *Implicit and explicit knowledge in second language learning, testing and teaching*. Bristol: Multilingual Matters, 2009. p. 65-93.

FERNÁNDEZ, E. M. *Bilingual sentence processing*. Amsterdam/Philadelphia: John Benjamins, 2003. <https://doi.org/10.1075/lald.29>

FERNÁNDEZ, E. M.; SOUZA, R. A. Walking Bilinguals Across Language Boundaries. In: HEREDIA, Roberto; ALTARRIBA, Jeannette. CIESLICKA, Anna. (Ed.). *Methods in bilingual reading comprehension research*. New York: Springer, 2016. p. 33-60.

FERNÁNDEZ, E. M.; SOUZA, R. A.; CARANDO, A. Bilingual innovations: Experimental evidence offers clues regarding the psycholinguistics of language change. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 20, no. 2, p. 251-268, 2017.

FORSTER, K.; FORSTER, J. DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers*, Springer Link, v. 35, p. 116-124, 2003.

FORSTER, K.; GUERRERA, C.; ELLIOT, L. The maze task: measuring forced incremental sentence processing time. *Behavior Research Methods*, Springer Link, v. 41, n. 1, p. 163-17, 2009.

GODFROID, A.; LOEWEN, S.; JUNG, S.; PARK, J-H.; GASS, S. Timed and untimed grammaticality judgments measure distinct types of knowledge – Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, Cambridge University Press, v. 37, n. 2, p. 269-297, 2015.

GOLDBERG, A.; JACKENDOFF, R. The English Resultative as a family of constructions. *Language*, Washington, Linguistic Society of America, v. 80, p. 523-567, 2004.

GROSJEAN, F. Bilingualism: A short introduction. In: GROSJEAN, François; PING, Li. (Ed.). *The Psycholinguistics of Bilingualism*. Oxford: Wiley-Blackwell, 2013. p. 5-25.

GUTIÉRREZ, X. The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition*, Cambridge University Press, v. 35, n. 3, p. 423-449, 2013.

HAYES, N.; BROADBENT, D. Two modes of learning for interactive tasks. *Cognition*, Elsevier, v. 28, n. 3, p. 249-276, 1988. [https://doi.org/10.1016/0010-0277\(88\)90015-7](https://doi.org/10.1016/0010-0277(88)90015-7)

JEFFERIES, E.; LAMBON RALPH, M.; BADDELEY, A. Automatic and controlled processing in sentence recall: the role of long-term and working memory. *Journal of Memory and Language*, Elsevier, v. 51, n. 4, p. 623-643, Nov. 2004. <https://doi.org/10.1016/j.jml.2004.07.005>

KIM, J.; NAM, H. Measures of implicit knowledge revisited – Processing modes, time pressure, and modalities. *Studies in Second Language Acquisition*, Cambridge University Press, p. 1-27, Feb. 2016. <https://doi.org/10.1017/S0272263115000510>

KRASHEN, S. The input hypothesis and its rivals. In: ELLIS, Nick (Ed.). *Implicit and Explicit Learning of Languages*. San Diego: Academic Press, 1994. p. 45-77.

LEVINSON, L. Finding arguments for pseudo-resultative predicates. *University of Pennsylvania Working Papers in Linguistics*, Philadelphia, PA, v. 13, n. 1, 2007.

LONG, M. Maturational constraints on child and adult SLA. In: GRANENA, Gisela; LONG, Mike. (Ed.). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam; Philadelphia: John Benjamins. 2013. p. 3-41. <https://doi.org/10.1075/llt.35.01lon>

LOWEN, S.; GONULAL, T. Exploratory factor analysis and principal component analysis. In: PLONSKY, Luke (Ed.). *Advancing quantitative methods in second language research*. London; New York: Routledge, 2015. p. 182-212.

NATION, P. *Teaching and learning vocabulary*. Boston: Heinle & Heinle, 1990.

OLIVEIRA, C. *A construção resultativa e sua representação por bilíngues do par linguístico português do Brasil e inglês*. 2013. 118f. Dissertação (Mestrado) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte. 2013.

OLIVEIRA, C. Uma investigação sobre a aquisição das regras de formação da construção resultativa por bilíngues de alta proficiência do par linguístico português do Brasil e inglês. *Letrônica*, Porto Alegre, Pontifícia Universidade Católica-RS, v. 7, n. 1, p. 66-85, 2014. <https://doi.org/10.15448/1984-4301.2014.1.16669>

OLIVEIRA, C. *The learnability, representation and processing of the resultative construction by Brazilian Portuguese-English bilinguals*. 2016. 200f. Tese (Doutorado) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte. 2016.

OLIVEIRA, C.; SOUZA, R. Uma exploração da aprendizibilidade da construção resultativa do inglês por bilíngues do par linguístico português do Brasil e inglês. *Confluência*, Rio de Janeiro, Linceu Literário Português, v. 43, n. 2, p. 242-260, 2012.

PAJAKA, B.; FINE, A.; KLEISCHMIDT, D.; JAEGER, T. F.; Learning additional languages as hierarchical probabilistic inference. *Language Learning*, Wiley Online Library, v. 66, n. 4, p. 900-944, 2016

PYLKKÄNNEN, L.; McELREE, B. The syntax-semantics interface: On-line composition of sentence meaning. In: TRAXLER, Matthew; GRENSBACHER, Marton. (Ed.). *The Handbook of Psycholinguistics*. 2nd ed. London; Burlington: Academic Press, 2006. p. 1-69.

RAMSCAR, M.; HENDRIX, P.; SHAOUL, C.; MILIN, P.; BAAYEN, H. The myth of cognitive decline: Non-linear dynamics of lifelong learning. *Topics in Cognitive Science*, Wiley Online Library, v. 6, p. 5-42, 2014.

REBER, A. Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, American Psychological Association, v. 118, n. 3, p. 219-235, 1989. <https://doi.org/10.1037/0096-3445.118.3.219>

ROBINSON, P.; MACKEY, A.; GASS, S.; SCHMIDT, R. Attention and awareness in second language acquisition. In: GASS, S.; MACKEY, A. (Ed.). *The Routledge handbook of second language acquisition*. London; New York: Routledge, 2014.

SCHMIDT, R. Attention. In: ROBINSON, P. (Ed.). *Cognition and second language instruction*. Cambridge: Cambridge University Press, 2001. <https://doi.org/10.1017/CBO9781139524780.003>

SCHMIDT, R. Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In: SCHMIDT, R. (Ed.). *Attention and Awareness in Foreign Language Learning*. Honolulu: University of Hawai'i Press, 1995.

SOARES-SILVA, J. *Exploring a vocabulary test and a judgment task as diagnoses of early and late bilinguals' L2 proficiency*. 2016. 150f. Tese (Doutorado) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte. 2016.

SOUZA, R. Memorização de sentenças com construção específica de uma das línguas de bilíngues. *Revista Letrônica*, Porto Alegre, Pontifícia Universidade Católica-RS, v. 7, n. 1, p. 86-112, 2014.

SOUZA, R.; OLIVEIRA, C. The learnability of the resultative construction in English L2: a comparative study of two forms of the Acceptability Judgment Task. *Revista da ABRALIN*, Associação Brasileira de Linguística, v. 8, n. 2, p. 375-410, 2014.

SOUZA, R.; OLIVEIRA, C.; SOARES-SILVA, J.; PENZIN, A.; SANTOS, A. Estudo sobre um parâmetro de tarefa e um parâmetro amostral para experimentos com julgamentos de aceitabilidade temporalizados. *Revista de Estudos da Linguagem*, Belo Horizonte, UFMG, v. 23, n. 1, p. 211-244, 2015.

SOUZA, R.; SOARES-SILVA, J. Exploring the measurement of vocabulary size to differentiate Brazilian Portuguese-English bilinguals' access to grammatical knowledge in the L2. *Linguística*, Rio de Janeiro, UFRJ, v. 11, p. 187-204, 2015.

SOUZA, R.; SOARES-SILVA, J.; SILVA, G.. Brazilian Portuguese-English bilingualism does not affect metalinguistic awareness of L1 constraints in two argument structure constructions. *Ilha do Desterro*, Florianópolis, UFSC, v. 69, n. 1, p. 17-32, 2016.

VAFAEE, P.; SUZUKI, Y.; KACHISNKE, I. Validating grammaticality judgment tests – Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, Cambridge University Press, p. 1-37, Jan. 2016.

WECHSLER, S. Resultatives and the problem of exceptions. In: LEE, Ik-Hwan *et al.* (Ed.). *Issues in English Linguistics* (Papers from the 1st World Congress of Scholars of English Linguistics). Seoul: South Korea; Hankook Munhwa; Hanyang University, Seoul, 2012. p. 119-131.

WECHSLER, S.; NOH, B. On resultative predicates and clauses: parallels between Korean and English. *Language Sciences*, v. 23, p. 391-423, 2001. [https://doi.org/10.1016/S0388-0001\(00\)00031-0](https://doi.org/10.1016/S0388-0001(00)00031-0)

WILSON, M. *Constructing measures: an item response modeling approach*. Mahwah, N.J; London: Lawrence Erlbaum Associates, 2005

WINTER, B.; REBER, A. Implicit learning and the acquisition of natural languages. In: ELLIS, Nick (Ed.). *Implicit and explicit learning of languages*. San Diego: Academic Press, 1994. p. 115-145.

Morphological priming resists language and modality switching in late Dutch-Brazilian Portuguese bilinguals

Priming morfológico resiste em face à alternância entre línguas e modalidades com bilíngues tardios de holandês e português brasileiro

Marije Soto

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
marijesoto@hotmail.com

Aline Gesualdi Manhães

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Rio de Janeiro, Rio de Janeiro / Brasil
alinegesualdi@gmail.com

Abstract: In this study, we aimed to investigate how late bilinguals of Dutch and Brazilian Portuguese (BP) engage in morphological processing in L2. Morphologically related and unrelated word pairs combined with cross-linguistic and cross-modal variations were presented in a lexical decision experiment in order to compare morphological priming effects in L2 and L1. We collected both reaction times and Event Related Potentials (ERPs). In L2, morphological priming effects were observed in lower amplitudes for the N400 component, but no effect was observed for RT measures. This can be explained under the assumption that the N400 component is more sensitive to lexical properties of roots (e.g. $\sqrt{\text{friend}}$) than to lexical properties of full word forms (e.g. *friendliness*). We consider that difficulties in L2 morphological processing are not necessarily due to a lack of parsing of morphemic units, but because of

the difficulty in the semantic interpretation of derivational morphemes in L2. This occurs via compositional mechanisms that are grammatically constrained, which are especially difficult to acquire in adult life. This is in line with what is suggested by the Superficial Processing Hypothesis (CLAHSEN *et al.*, 2010). Moreover cross-linguistic priming showed that although morphological priming occurred for switching from L2 to L1, it caused the opposite effect for switching from L1 to L2. These results confirm predictions made by the Revised Hierarchical Model (KROLL *et al.*, 2010), which explains this type of data in terms of the relative weakness of concept to word form mapping in L2 as compared to word form to concept mapping.

Keywords: switching; bilingual word recognition; N400.

Resumo: Neste estudo, pretende-se investigar o processamento morfológico em bilíngues tardios de holandês (L1) e português brasileiro (L2). Pares de palavras morfológicamente relacionados e pares não relacionados foram apresentados em uma tarefa de decisão lexical com alternância de línguas e de modalidade entre prime e alvo a fim de comparar efeitos de *priming* morfológico na L1 e na L2. Coletamos tempos de resposta (TRs) e Potenciais Relacionados a Eventos (ERPs). Na L2, efeitos de *priming* morfológico foram observados nas amplitudes reduzidas do componente do N400, mas nenhum efeito foi observado para os TRs. Isso pode ser explicado, se considerar que o N400 é mais sensível às propriedades lexicais das raízes (ex. *√amigo*) do que às propriedades lexicais das palavras inteiras (ex. *amigável*). As dificuldades no processamento morfológico na L1 não são necessariamente devidas a uma falha na segmentação dos morfemas, mas devido à dificuldade para interpretar os morfemas derivacionais semanticamente. Isso depende de mecanismos composicionais que são guiadas pelas regras morfológicas que são especialmente difíceis para serem adquiridas durante a vida adulta. Essa conclusão está de acordo com a Hipótese de Processamento Superficial (CLAHSEN *et al.*, 2010). Ainda, a alternância entre línguas evidenciou que, embora houvesse *priming* morfológico para a direção de L2 para L1, para a direção de L1 para L2 o efeito oposto ocorreu.

Esses resultados confirmam as previsões lançadas pelo Modelo Revisado Hierárquico (KROLL *et al.*, 2010), que explica esse tipo de dado em termos da fragilidade relativa na L2 do mapeamento do significado da palavra para a sua forma lexical comparada ao mapeamento da forma lexical para seu significado.

Palavras-chave: acesso lexical bilíngue; alternância; N400.

Recebido em 13 de dezembro de 2016

Aprovado em 7 de maio de 2017

1 Introduction

The agility with which bilinguals alternate freely between lexical items in different languages is truly an amazing feat of human cognition. Yet it is part of everyday life for the majority of people in the world. A conservative estimate of about 50% of the world population (BHATIA; WILLIAMS, 2005) operates easily in both monolingual and bilingual mode¹ and engages in code switching² on a daily basis. This has led many researchers to raise questions about how the bilingual lexicon is

¹ According to Grosjean (1989), the mode in which bilinguals activate L1 and L2 refers to whether both languages are equally activated (bilingual mode), or asymmetrically activated (monolingual mode). The mode in which the bilingual operates depends on the information he or she receives in the form of expectancy, pragmatic context, or task instructions.

² Code Switching is a typical phenomenon among bilingual speakers. It consists of inserting excerpts of speech in one language, in a conversation structured in another language. For a long time, this phenomenon was understood as a way to compensate for some deficiency in proficiency in one of the dominant languages of the bilingual speaker. The hypothesis was that when bilingual speakers engage in code switching, it is to compensate for some aspect they do not master in one of their native languages by inserting words of the other, a phenomenon known as semilingualism (LIPSKY., 1978; HEREDIA; BROWN, 2004). Currently, it is believed that Code Switching involves pragmatic competence of bilinguals which enables them to recognize bilingual conversational partners of the same languages, thus sharing in common linguistic competence and cultural identity (MEISEL, 1994).

organized, the manner in which lexical representations are activated, and the nature of the processing mechanisms operating on the comprehension and production of lexical items in monolingual and bilingual modes. Essential to these questions is also how factors such as age of acquisition, exposure, cultural identity, among others, influence these aspects.

In this study, we focus on whether and how late bilinguals engage in morphological processing in L2. With regards to this, the main issues under debate are the following: (i) is there any morphological processing in L2 for late bilinguals, and if so, is it comparable to L1 morphological processing?; (ii) are (morphological) processing strategies dependent on L1 characteristics? (for example, if L1 is morphologically 'rich' in comparison to L2, does this lead to deeper L2 morphological processing?); (iii) does L2 morphology map onto meaning as readily as L1 morphology does, or is it mostly mediated by lexical representations in L1? (BRAESBART; DIJKSTRA, 2006; PORTIN *et al.*, 2007; LEMHÖFER *et al.*, 2008; DIEPENDAELE *et al.*, 2011). We approached these questions by investigating morphological priming³ in late bilinguals who are native speakers of Dutch (their dominant language), having learned Brazilian Portuguese as adults (>21y) while living in Brazil. We adhere to Grosjean's definition of bilinguals. He defines bilinguals as individuals who use two languages in their daily life (GROSJEAN, 1994). To tap into the processing mechanisms of this group of bilinguals we collected both neurophysiologic and behavioral data, using Event Related Potential (ERP)⁴ and priming methodologies. There are very few

³ In a word priming paradigm, different types of relations between prime (1st word) and target (2nd word) are manipulated, such as: semantic association (*apple-pear*), morphological relation (*teach-teacher*), phonologic similarity (*peach-pea*), repetition (*peach-peach*), to name a few. The rationale behind this paradigm is that the activation of the second word is affected (facilitated or hindered) by the activation of the previously presented word (the prime). Depending on the research hypothesis, a specific kind of relation can be manipulated in order to assess its involvement in word representation and processing.

⁴ Event Related brain Potentials (ERPs) are extracted from continuous Electroencephalogram (EEG) recordings in which stimulus presentation onset is marked so that the signal can be segmented to study neurophysiological responses to stimuli. Segments are added and averaged per experimental condition so that they may be compared for experimental condition effects. This is measured by comparing amplitudes (in voltage) and latencies (moment of maximum peak in ms) in a given time interval. An N400 component is thus a negative amplitude peak approx. 400ms after stimulus onset, a P300 effect is a positive peak approx. 300ms after stimulus onset.

studies on morphological priming in bilinguals using ERP methodology, and none of them comparing Dutch-Brazilian Portuguese bilinguals. Therefore, as a pioneer contribution to the field, this study might raise more questions than it answers, as there are as of yet not many data to compare with in literature.

In order to address the specific question of morphological processing in L2, participants had to engage in a lexical decision task, presenting morphologically related pairs (ex. *friend-friendly*) and unrelated pairs (ex. *pencil – garden*) as the control condition in both L1 and L2. Facilitation of lexical access by morphological priming as compared to a control baseline has been attested both in monolinguals and bilinguals (DIJKSTRA *et al.*, 2005; PORTIN *et al.*, 2007; SILVA; CLAHSN, 2008; DOMINGUEZ; SEGUÍ; CUETOS, 2002), and -as an experimental effect- can be isolated from phonological, orthographical and semantic priming effects (DE GROOT; DELMAAR; LUPKER, 2000; DIJKSTRA *et al.*, 1999). Target words that are morphologically related to their primes (e.g. *teach-teacher*) are recognized faster compared to targets in orthographically or phonologically related pairs (e.g. *freeze-free*), and semantically related pairs (e.g. *horse-farm*) (DIEPENDAELE *et al.*, 2011). This specific facilitation effect indicates that morphological information is an integral and influential part in the architecture of the lexicon, both for monolinguals and bilinguals. However, the manner in which morphology operates in bilingual comprehension, in mono or bilingual mode, is still a matter of debate.

Although morphological priming effects have been shown in several studies with bilinguals, researchers are not conclusive in the interpretation of the results. A study by Diependaele *et al.* (2011) showed that Spanish-English and Dutch-English bilinguals, who had started learning English from the ages of 18 and 12, respectively, displayed similar reaction times as English monolinguals in a masked morphological priming experiment in which participants had to perform a lexical decision task in English(L2). Some authors claim data such as these suggest that bilinguals largely adopt the same processing strategies in L2 as monolingual native speakers of that language (LEMHÖFER *et al.*, 2008). If this is indeed the case, it means that morphological processing in L2 involves parsing words for morphemic units and that recognition of morphologically complex words is guided by rule based automatic processing, analogue to L1 processing. Contrary to this idea,

other authors suggest that bilinguals rely more heavily on whole-word processing in their second language than in their first language. In this case a stored representation of a whole-word is accessed, while morphological composition within words remains opaque (ULLMAN, 2004, 2005; CLAHSEN *et al.*, 2010). Data that point in this direction are presented by Hahne *et al.* (2006). Their study showed that late bilinguals use morphological decomposition strategies to process morphological derivation with suffixes, such as in *bitter-bitterness*, but not to process inflectional morphology, either for regular or irregular verbs.⁵ Some authors take this to be evidence that derivational and inflectional processes have differential cognitive status, given that the latter depends on interpretation of functional categories, such as tense features, for which L2 grammar representation may not contain the same level of detailing as compared to L1 grammar (ULLMAN, 2004, 2005; HAHNE *et al.*, 2006; CLAHSEN *et al.*, 2010).

In this view, lack of detailing in L2 grammar occurs when L2 is learned after childhood or, for some authors, after puberty. For Ullman (2004, 2005) this idea is based on his declarative/procedural model. In this model, word recognition can occur via direct lexical retrieval or via compositional mechanisms that are grammatically constrained. The former relies on declarative memory, subserved by the left temporal lobe, and the latter on procedural memory, which engages the left frontal cortex and basal ganglia. Rule based automatic procedures are acquired early in life, and, thus, do not support language learning after childhood. This implies that late bilinguals are more dependent on declarative memory for L2 representations, and, therefore, are biased

⁵Most commonly morphemes are understood to be of two categories: (i) derivational; or (ii) inflectional. Derivational morphemes add to the meaning of a root, affecting both category and meaning (ex. *to judge* (v)-*judgment* (n)), and morphological rules supposedly determine which combinations between roots and morphemes are grammatical (ex. *-ment* combines with a specific group of *verbs*). Inflectional morphemes do not change the category of the word they combine with, but indicate syntactic or semantic relations between constituents of sentences. For example, the present tense morpheme *-s* indicates agreement between a verb and its 3rd person subject. However, there are authors that contest the cognitive relevance of the derivational vs. inflectional distinction, given that there are some morphemes that share characteristics of both types (e.g. a participle morpheme can both nominalize and inflect in BP) (MARANTZ, 1997)

for whole-word representations. Some authors associate this supposed inability for deep grammatical processing in late bilinguals as a logical consequence of cognitive-biological limitations for language acquisition beyond the critical period (HAHNE; FRIEDERICI, 2001). Other authors stress the possibility of the development of proficiency over time, even for late learners, which greatly attenuate effects of so-called superficial processing (HAHNE *et al.*, 2006; KOTZ *et al.*, 2004).

The reduced ability for acquiring rule based automatic procedures for late bilinguals supposedly leads to an L2 grammar that is morphologically underspecified, consequently only enabling 'superficial' processing (Superficial Processing Hypothesis, CLAHSEN; FELSER, 2006; CLAHSEN *et al.*, 2010). Thus, the functioning of the morphological parser in bilinguals is expected to yield overall slower and less automatized processing, especially for late L2 learners, who supposedly rely more heavily on memory and attention for L2 processing than for L1 processing (ULLMAN, 2004; 2005). Mostly this rationale is built on evidence showing bilinguals process derivational morphology differently from inflectional morphology. Nonetheless, criticism on this rationale is that morphological priming results in bilinguals mostly resemble monolingual data, which also tend to show stronger morphemic priming with derivations than with inflections (FELDMAN, 1994; SCHRIEFERS; FRIEDERICI; GRAETZ, 1992; SILVA; CLAHSEN, 2008). That is why, in this study we chose to focus only on the stimuli with derivational suffixes.

Based on data from the literature, an effect for morphological priming for late bilinguals in both L1 and L2 is expected, especially if the stimuli only present derivational morphemes. On the other hand, morphological processing might be relatively superficial for L2 as compared to L1, due to late acquisition of L2. If this affects parsing across the board, RTs might be reduced for targets in L2.

To ensure that potential morphological priming in L2 is not merely due to an orthographic form repetition effect, participants were presented with trials in which modality was varied. Primes were presented in either auditory or written form, while targets were always written. In lexical decision tasks, auditory stimuli generate reduced amplitudes for ERPs and faster RTs, reflecting ease of processing as compared to written stimuli (STOCKAL; MARANTZ, 2006). Yet in this study, primes will be varied for modality, while all targets will be written. In the case of

cross-modal priming pairs, a switching cost is generated, as evidenced by slower RTs and slower N400 effects in comparison to unimodal priming (HOLCOMB *et al.*, 2005; KIYONAGA *et al.*, 2007).

Proficiency levels in late bilinguals affect both RT measures and ERP data. ERP studies show that proficient L2 speakers tend to present similar neurophysiologic responses as native speakers of that language (KOTZ *et al.*, 2008). The N400 component can appear in both less and more proficient L2 speakers. This ERP signature is generally associated with semantic processing, and it also sensitive to morphological processing. Facilitation of access to target words as a result of morphological or semantic priming yields lower amplitudes 400ms after word presentation onset as compared to targets of unrelated word pairs. Generally, the lower amplitudes are interpreted as reflecting a relative ease in processing (OSTERHOUT *et al.*, 1997; MUNTE *et al.*, 1999; KUTAS; FEDERMEIER, 2000; PYLKKÄNEN *et al.*, 2000; 2003; DOM *et al.*, 2004; MARANTZ, 2005; FRANÇA *et al.*, 2008; LAU *et al.*, 2008; SOTO, 2010; GARCIA *et al.*, 2012). Differences in proficiency may correlate with subtle modulations in latency and amplitude of the N400 component.

Even speakers with relatively low proficiency can present N400 responses to semantic priming. However, very few studies have looked into bilingual morphological priming with ERP methodologies. Components that supposedly reflect (morpho)syntactic processing, such as ELAN and P600, generally only appear in very proficient speakers (KOTZ, 2001; DIJKSTRA *et al.*, 2005; HAHNE; FRIEDERICI, 2001). Nonetheless, P600 waveforms can appear in proficient late bilinguals when processing L2, and they may resemble those of native speakers of that same language. Hahne *et al.* (2006) presented participants with inflected verbs (past tense) and nouns (plural) in L2 in a syntactic violation paradigm, which yielded a P600 effect, and not an earlier ELAN component. The authors took this to as a confirmation of the aforementioned claim that bilinguals process inflectional morphology different from derivational morphology. Factors that influence resemblance to native-like responses for N400 and P600 components are proficiency (TANNER *et al.*, 2016), age of acquisition (KOTZ *et al.*, 2008) and the degree of similarity of the tested syntactic property between L1 and L2 (DÍAZ *et al.* 2016). In this study, we would expect reduced N400 amplitudes for targets in morphologically related pairs as compared to targets in unrelated word pairs. The degree of this

morphological priming effect on the N400 components might reflect the proficiency and stage of acquisition of the group of late bilinguals.

Another hypothesis, put forth by Díaz *et al.* (2016) as well as Portin *et al.* (2007), is that language specific traits might influence both RT and ERP measures. Portin *et al.* (2007) discuss whether language specific morphological traits influence processing strategies. They suggest that native speakers of languages that are known to be morphologically complex, such as Finnish, which boasts 200 possible declinations for nouns alone, apply similar processing strategies on languages that display relatively less rich derivational and inflectional morphology, such as English. The authors claim that bilingual native speakers of Finnish will display a tendency to use morphological decomposition for L2 word recognition because their native language has ‘sculpted’ processing mechanisms. This view, stemming from a weaker version of the Linguistic Relativity thesis (WHORF, 1956; BORODITSKY, 2003) would predict that perhaps Dutch L1 speakers are biased for monomorphemic and monosyllabic words (e.g. *tuin*, (“garden”)) and productive compounding processes (e.g. *tuin-man*, (“garden-man -> gardener”)). Derivational morphemes are productive in Dutch (e.g. *vriend – vriendelijkheid*, (“friend – friendliness”)), but compounding is far more common in word formation. On the other hand, compounding as a word forming process is rare in BP, which is why all targets (in L1 as well as L2) in this study are words with two or more derivational morphemes. An L1 bias in this group of Dutch-BP speakers might make them less likely to engage in processing mechanisms that are inherent to morphologically rich languages such as Brazilian Portuguese (BP). If this is case, than contrary to our initial prediction, morphological priming effects in L2(BP) might be weak or even absent, meaning that RTs and N400 will show little difference with those in response to unrelated pairs.

Another way of investigating morphological processing in L2 is to test whether L2 morphemes map onto meaning just as readily as L1 morphemes do. In order to examine this, we included cross-linguistic morphological priming in both directions: prime L1-target L2, and prime L2-target L1. There is now a large amount of data in the literature on bilingual word recognition that steadily points to bilingual lexical access as being non-selective for either L1 or L2. This means that in the first stages of auditory and visual word recognition, in principle, word candidates from the different languages may compete (JACKET;

FRENCH, 2002; BRYSSBAERT; DIJKSTRA, 2006; KROLL *et al.*, 2010). This notion is reflected by the various hypotheses of how the bilingual lexicon is organized. Traditional models for bilingual word recognition, such as the one proposed by Paradis (1997), propose language-specific storage for word forms (orthographic and phonological forms with their syntactic properties) and word meanings, with a conceptual non-linguistic level shared by both languages.

The idea of a separation between language-dependent lexical meaning storage and a conceptual level shared by two languages is still present in more recent models, such as the Revised Hierarchical Model (RHM) (DE GROOT, 1998; KROLL; DE GROOT, 1997; KROLL *et al.*, 2010). Yet the focus has shifted to the notion of two lexicons that are functionally separate, rather than a physical or organizational separation for the storage of bilingual lexicons. Moreover, data indicate that parallel activation to lexical and sublexical information between languages occur (KROLL *et al.*, 2010). The question is thus if meaning activated by a word in L1 (ex. *vriend* (“friend”)) can prime a root in a morphologically more complex word in L2 (ex. *√amigo+(a)vel* (“friend + ly”)), provided that the root is a translation candidate of the word in L1, and vice versa. The RHM, although more intended as a model for word prediction than word recognition, foresees that L1 and L2 can both map onto conceptual representations, but that links between L2 word forms and concepts are weaker. More specifically, the RHM hypothesizes that in L2, the strength of links in the direction of concept to word form is weaker than that for the opposite direction, from word form to concept. This may cause asymmetric effects in the sense that, for late bilinguals, access from words to concepts may be accomplished easily, while access from concepts to words may demand more effort. Under this assumption, RHM would predict facilitation from L2 to L1 (L2L1: *amigo-vriendelijk* (“friend-friendly”)), given that the L2 word form *amigo* readily maps onto a conceptual representation, which has rich links to L1, thus, easily priming morphologically related candidates. Additionally, there might be a parallel link connecting the word *amigo* directly to the word form of a translation pair (e.g. *vriend*). Yet, language switching in the other direction (L1L2: *vriend-amigável* (“friend-friendly”)) would be expected to make target recognition more difficult due to the fact that the L1 word easily maps on to a conceptual representation, which might be widely connected to an array of conceptual representations that link to

other lexical representations in L1. Not only would these compete with any upcoming word in L2, but mapping from concept to form might be more difficult in L2. Added to this asymmetric effect is that even if the conceptual features are shared, access to these features via word form, might not be equivalent for L1 and L2, especially if bilinguals are less proficient, for example (KROLL *et al.*, 2010). Therefore, if the RHM is correct in its predictions, for this study, we might expect facilitation to hold when there is a switch between prime and target language, depending on the switching condition. Facilitation is expected for L2 to L1, whereas slow processing is expected for L1 to L2.

However, not all studies have shown evidence of asymmetric cost, especially not in word recognition tasks. For word recognition, there is not necessarily a need to select among competitors, especially if the task is to merely acknowledge the word as being a potential word in either L1 or L2, such as is the case in this experiment. Also, in this study, bilinguals are consciously triggered into bilingual mode because monolingual priming pairs (L1L1 or L2L2) are alternated with bilingual pairs (L1L2 or L2L1) and they are told to recognize words in either L1 or L2. Therefore, the cognitive effort spent on inhibiting L1 competitors may be reduced (BRYSSAERT; DIJKSTRA, 2006). There are also other factors that may reduce asymmetric switching effects. The strengthening of links between L2 word forms and conceptual representations may evolve as bilinguals gain more proficiency.

ERP data for morphological processing in monolinguals present evidence that word recognition is compositional, given that N400 components are sensitive to lexical properties of roots (e.g. $\sqrt{\text{friend}}$) more so than lexical properties of the full word form (*friendliness*). This suggests that N400 responses may reflect access to the stored meaning representation as activated by a given lexical root, and that subsequent processing of derivational suffixes is of relatively low cost, not reflected by N400 amplitudes, perhaps since it only adds highly predictable and systematized meaning to an already accessed meaning (e.g. $\sqrt{\text{friend}}$ + *ly* (attribute) + *-ness* (state of)) (PYLKKÄNEN *et al.*, 2000, 2002, 2003; FRANÇA, *et al.*, 2008).

In the case of cross-linguistic morphological priming, the prime and target are related by way of translation pairs that supposedly map on to shared conceptual features. For example, if the prime is *vriend* (L1 (“friend”)), it may map on to *amigo* (L2 (“friend”)), which is the

lexical root of the morphologically more complex target *amigável* (L2 (“friendly”)). An N400 component measured during the presentation of a target in L2 probably reflects facilitation due to the sustained activation of the meaning representation as accessed by the prime *vriend*. If we assume that the lexical root in L2 and translation pair of the prime *amigo* is easily identified within the morphologically complex target by morphological parsing, its mapping onto an already activated meaning should yield facilitation effects. However the posterior process of parsing (in this example: *amig-(a)vel*) and additional processing of derivational suffixes might be more costly in L2 than it is in L1. This might be caused by less efficient morphological parsing in L2 or by difficulty mapping the suffixes onto meaning, as well as a reduced richness in the representation of morphologically related word forms in the L2 lexicon. All of these factors might make it more difficult to recognize a morphologically complex target word in L2, even if it is primed by a morphologically related word in L1. Thus, while N400 components may show facilitation effects, reaction times may not due to these additional difficulties in the case of L1 to L2 switching. In the ERP literature opposite effects between behavioral results and N400 components are not uncommon (HAHNE; FRIEDERICI, 2001; HAHNE *et al.*, 2006, KOTZ, 2001; KOTZ; ELSTON-GÜTTLER, 2004; WEBER-FOX; NEVILLE, 1996), since these measures may reflect different stages in word recognition, or be affected differently by experimental task requirements.

According to the predictions of the RHM, as well as the Superficial Processing Hypothesis, an asymmetric switch effect is expected, but for different reasons. The RHM predicts that relatively weaker links between word forms and concepts in L2 will result in slower RTs for morphological priming, especially when the direction of the switch is from L1 to L2. The Superficial Processing Hypothesis foresees overall less efficient morphological processing for L2, which may also lead to an asymmetric switch effect from L1 to L2, but it also predicts no or reduced morphological priming for non-switch priming in L2.

2 The experiment

In this study we presented morphologically related and unrelated word pairs in order to compare morphological priming effects in L2 and L1. Facilitation for morphologically related targets relative to unrelated target controls in L2 would be evidence of effective morphological processing in late bilinguals of Dutch and Brazilian Portuguese (BP). Participants engaged in a lexical decision task, for which they had to decide whether targets were words either in Dutch or in BP. In order to present an equal number of trials with possible yes or no answers, we added trials with Hungarian targets, a language none of the participants had been exposed to before. We collected EEG measures and reaction times (RTs) from 18 proficient late bilinguals. Faster RTs and lower mean amplitudes of the N400 component reveal facilitation effects for morphological priming as compared to unrelated prime and target pairs.

Besides, non-switch trials (L1 to L1 and L2 to L2), participants were also presented with switch trials (L2 to L1 and L1 to L2). With the switch condition, we aimed to test whether L2 morphemes map onto meaning just as readily as L1 morphemes do. The lexical root contained in a morphologically complex target word may be primed by a prime if it is a translation candidate (e.g. prime: *vriend* (L1 (“friend”)) – target: *√amig- (a)vel/ amigável*, (“friend-ly”)). The Revised Hierarchical Model foresees that the switch direction from L1 to L2 is more costly because in L2 the links from concept to word form might be weaker than those mapping word forms to concepts. The Superficial processing hypothesis (CLAHSEN; FELSER, 2006; CLAHSEN *et al.*, 2010) foresees overall less efficient morphological processing for L2, which may also lead to an asymmetric switch effect from L1 to L2, but it also predicts no or reduced morphological priming for non-switch priming in L2. Asymmetric priming effects as well as slower overall processing for L2 in general may be reflected by relatively slower RTs as compared to L1 targets, or absence of any morphological priming. However, N400 amplitudes may yield evidence for facilitation, even when RTs do not, if we assume that N400 components sensitive to lexical properties of roots (e.g. *√friend*) than to lexical properties of full word forms (e.g. *friendliness*).

Finally, to ensure that the morphological priming effects are not merely a case of an orthographic repetition effect, we added a cross-modal variable, presenting trials with auditory primes followed by written targets, as well as trials for which both prime and target were written. For both types of trials, morphological priming effects may be expected.

2. 1 Materials and Methods

The four variables in our study are: (i) relation type: morphological vs. unrelated; (ii) switch: switch vs. non-switch; (iii) target language: L1(Dutch) vs. L2(BP); (iv) modality: auditory vs. written. These combinations yielded 16 conditions. In Table 1, conditions and sample stimuli are presented except for those including Hungarian (H) targets. There were four conditions with Hungarian targets: L1H and L2H, both with auditory and written primes. Modality conditions are not specified in Table 1, but of all of the 8 conditions were presented with auditory primes and written primes, reaching a total of 16 conditions. Each condition was represented by 8 items, and an equal number of pairs with Hungarian targets were added for the purposes of the lexical decision task. In total, 256 pairs were presented to each participant (for the complete list, see Appendix 1).

The auditory stimuli were recorded by native speakers for L1 and L2, both speakers were female and matched for pitch. We controlled words for frequency over all conditions by using a hit count in Google,⁶ and equally distributed the number of items for all the grammatical classes of the primes(L1: noun=83; adj=25; verb=20 and L2: noun=82; adj=25; verb=21), and targets (L1: noun=30; adj=32; verb=2, and L2: noun=34; adj=25; verb=3; adverb=2). We distributed a controlled variety of suffix types over all conditions, presenting at least 2 items for each suffix type (L1: 9 different suffixes and L2: 14 different suffixes) (see Appendix 2 for a complete list). Also, we controlled for number of phonemes, presenting a range of 4-6 for primes, and 6-8 for targets. In terms of morphological complexity, primes were always composed of a root and a category morpheme (e.g. *amig* + *o*, (“friend”)), and the target was always a morphologically complex word with 2 or more morphemes (e.g. *amig* + *a* + *vel*, (“friendly”) or *vriend* + (*e*) *lijk* + *heid*, (“friendliness”)). Stimuli were pseudo-randomized and divided into two

⁶In the absence of large corpora for Brazilian Portuguese, we chose to run a Google search for each word in the stimuli set. Frequency is measured by the number of *Google hits* for *.br* sites on a given day. It is perhaps not an ideal but an acceptable indication of surface frequency in the absence of large accessible corpus with the appropriate information codification. *Google hits* filter out infrequent words such as *drupa* (‘blossom’) with 24.500 hits vs. *flor* (‘flower’) with 10.400.000 hits or *coturno* (lumberjack) 955.000 hits vs. *bota* (‘boot’) with 5.360.000 hits

lists. Eight participants from a previous Dutch-BP bilingual study who did not participate in this experiment were asked to judge all Portuguese words for familiarity. They were told to mark words as either *known*, *unknown*, or *I'm not sure what this word means*. None of the items were rejected unanimously; however, out of all the 176 Portuguese words, there were nine items which were marked as unknown by three volunteers. Notwithstanding, given that only one of these nine items was marked twice, we did not exclude any items.

TABLE 1 – Experimental design: experimental design and prime-target examples

switch:	prime-target relation:	target L:	
		Dutch (L1)	Brazilian Portuguese (L2)
switch	morph.related	<i>jogo</i> – <i>SPELER</i> ("game-GAMER")	<i>sluiten</i> – <i>FECHAMENTO</i> ("close-CLOSURE")
	unrelated	<i>rotina</i> – <i>VRIENDIN</i> ("routine-GIRLFRIEND")	<i>winnen</i> – <i>GRÁVIDA</i> ("win-PREGNANT")
Non-switch	morph.related	<i>moeder</i> – <i>MOEDERLIJK</i> ("mother-MOTHERLY")	<i>casar</i> – <i>CASAMENTO</i> ("wed-WEDDING")
	unrelated	<i>vraag</i> – <i>PLEZIER</i> ("question-PLEASURE")	<i>burro</i> – <i>SEGURO</i> ("dumb-INSURANCE")

2.2 Participants

In this study we take bilinguals to be individuals who make use of two languages in their daily life (GROSJEAN, 1994). Eighteen participants took part in the experiment, but 2 datasets were excluded due to artifacts and excessive noise on the EEG signals. One participant was excluded from the analysis based on proficiency scores. Thus, 15 datasets were included in the analysis. All participants had lived in Brazil for an average of 6,8 years (ranging from 4 to 18 years) and had thus been exposed to BP for the extent of that period. Participants varied in

age from 25 to 48 years. None of the participants had been exposed to PB under the age of 21.

All participants reported that L1(Dutch) was their dominant language. L2 Proficiency varied somewhat between participants, but before partaking in the experiment, participants were asked to describe some images in a *Monica e sua turma* comic book, the results of which were recorded and post-hoc assessed by 3 independent judges (native speakers of L2, and graduate students of the Language Faculty at the Federal University of Rio de Janeiro). The judges used a three-point scale to assess participants on pronunciation, vocabulary, and fluency, and a five-point scale for structure. Most participants scored between 2 and 3 points for pronunciation and vocabulary and around 4 points for structure (see Appendix 4 for all results). One participant was excluded based on proficiency scores, which were far below the grades of the other participants, achieving only 50% for fluency and vocabulary.

Some biographic data were collected also (see Appendix 3). All participants had started learning Portuguese after the age of 21 and additionally spoke another foreign language (English, French, German, and Spanish were the languages mentioned), a possible confound which is hard to avoid when dealing with Dutch bilinguals. All participants held at least bachelor degrees, except for two participants who finished preparatory scientific education. All participants were right-handed and had normal or corrected to normal vision. Also, none of the participants reported having any hearing impairments. All participants signed a form of consent which had been pre-approved by the Ethical Committee of the University Hospital of the Federal University of Rio de Janeiro.

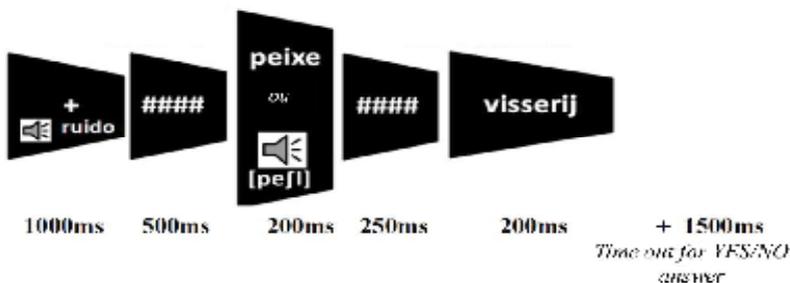
2.3 Procedure

Participants were placed at a 90cm distance in front of a 15 inch screen on which words were presented in Times New Roman font, size 25, on a black screen (see Figure 1 for the presentation protocol) using Presentation software version 12.2 (developed by Neurobehavioral Systems). The lexical decision task was first explained, and participants were instructed to use a button box to respond, pressing a green button for YES and a red button for NO, only judging the target. Left and right positions of buttons were alternated between participants. There was a training session of 15 trials, which was repeated if participants did not

understand the task correctly. During the experiment, there were four breaks during which participants could continue at their own leisure. Total time of recording, including electrode placement and proficiency testing, was approximately 1.5 hours.

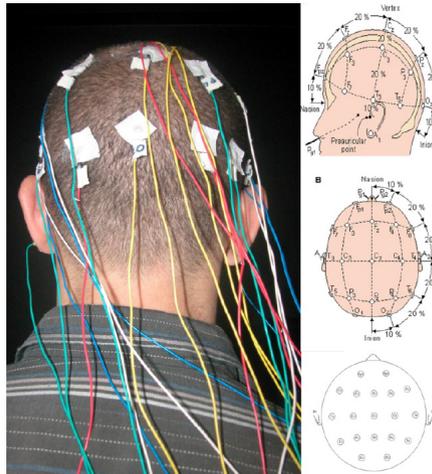
The presentation protocol of the stimuli is shown in Figure 2. A low pitched tone marked the beginning of each trial, accompanied by a fixation cross. Primes and targets were each presented for 200ms separated by masking for 250ms. Primes alternated for either sound or written form. After target presentation, there was a 1500ms time out period during which participants carried out the lexical decision task.

FIGURE 1 – Presentation protocol: presentation protocol for the priming paradigm as presented on the Presentation platform. The last screen presents the target for which the participant has to perform the lexical decision task



Before starting the experiment, 21 electrodes were placed on the participant’s head according to the international 10–20 set up (see Figure 2) (NIEDERMEYER; SILVA, 1982). Silver cup electrodes were fixed to the scalp with electrolyte paste (Elefix by Nihon Kohden). Two electrodes placed on the left and right earlobes served as reference. Impedance was kept below 10 kΩ throughout the experiment. The EEG equipment used was BrainNet BNT 36, and EEG signals were filtered during the recording session (100Hz, and 0,1Hz), sampling frequency was 200Hz.

FIGURE 2 – Electrode set up: The photo shows a participant from the study to illustrate electrode placement according to standard 10-20 configuration



3 Data processing and analysis

Reaction times were recorded using the Presentation platform, and later on analyzed for normal distribution (Mauchly's Test of Sphericity). A cut off of 95% accuracy rate was determined a priori, but since all the participants achieved 95% accuracy or over, none of the datasets were rejected. All trials were included in the analysis, except for outliers with a Z-score higher than 3 or lower than -3. RTs were analyzed using a four-way ANOVA Repeated Measures analysis in SPSS software, version 20. RTs were analyzed for all experimental variables: (i) switch; (ii) relation type; (iii) target language; and (iv) modality. RTs in response to Hungarian were not included in the analysis because they were not considered relevant to the issues under investigation.

Raw EEG signals were digitalized and processed in BrainVision Analyzer (BVA) software, version 2. Due to excessively noisy EEG signals, two datasets were excluded from the analysis. From the processed ERP wave forms (filtered at 30Hz, baseline correction 200ms) we extracted mean amplitudes over a predetermined interval of 350-450ms after word presentation onset. The mean amplitude analysis requires that an a priori time interval is chosen from which all data points are taken into account to calculate a mean. In this experiment, sampling frequency was 200Hz, which means that within the 350-450ms interval 20 data points

yield one mean amplitude value (in μV). Extracting mean amplitudes is considered a more reliable method than extracting only one maximum amplitude, based on the highest amplitude within a given time interval. This is because, in the averaging of the individual segments that form the final ERP wave forms, distorting can occur due to variability among items and subjects. Therefore, it is more relevant to look at all the data points within an interval as opposed to only 1 (maximum) data point (LUCK, 2005).

Based on anatomic region, electrodes with similar anatomic lobe reference were pooled into one Region of Interest (ROI) by the BVA pooling function, yielding the following ROIs: Frontal (F3,F4,F7,F8,Fz); Central (C3, C4,Cz); Temporal (T3,T4,T5,T6), Parietal (P3,P4,Pz), and Occipital (O1,O2,Oz). Mean amplitude values were exported, and analyzed in SPSS (version 20), applying a five-way ANOVA Repeated Measures analysis. Mean amplitudes were analyzed for all experimental variables: (i) switch; (ii) relation type; (iii) target language; and (iv) modality, including (v) ROI (Frontal, Central, Temporal, Parietal, and Occipital) as a variable. For all ANOVA analyses, p values of <0.05 were considered statistically significant.

4 Results

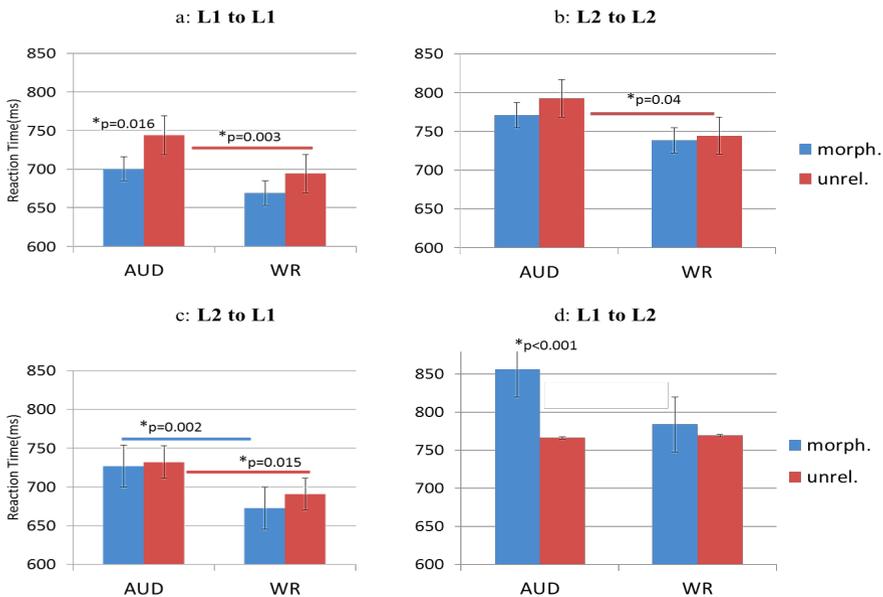
4.1 Behavioral Data: results

RTs were analyzed using a four-way ANOVA Repeated Measures analysis for the variables: (i) switch: switch x non-switch; (ii) relation type: morphological vs. unrelated; (iii) target language: L1 vs. L2; and (iv) modality: auditory X written. Main effects were found for all variables except relation type ($F(1,118)=0.082$, $p=0.776$). The main effect for the target language variable ($F(1,118)=150.48$, $p<0.001$) revealed that overall L1 rendered fastest RTs: both for non-switch 703.66ms(163.21ms) and for switch 708.38ms(169.71ms) conditions. Post-hoc pairwise comparisons (LSD) showed there was no significant difference between the two conditions ($p=0.60$). Overall L2 as a target language yielded slower RTs: for non-switch 766.12ms(7.78ms) and for switch 793.61ms(7.78ms). There was a significant difference between the conditions with a p-value of 0,008. Differences between target languages L1 and L2, irrespective of the switching condition, were always significant with p-values of $<0,001$. The main effects for the switch variable ($F(1, 118)=7.53$, $p=0.007$) revealed overall slower RTs for the switching condition (non-switch: 734.99ms(5.89ms) vs. switch: 750.99ms(5.98ms)). However, as the interaction between the effects of

Target language and switch variables indicates (Target L. vs. Switch: $F(1,118)=5.33, p=0.023$), switching from L1 to L2 yielded the slowest RTs (793.61ms(7.78ms)). The main effect for modality ($F(1,118)=28.64, p<0.001$) was reflected by overall significantly slower RTs for targets followed by auditory primes as compared to written primes (AUD: 766.15ms(5.41ms) vs. WR: 721.05ms(5.14ms), $p<0.001$).

Various interactions were also found: switch X target language: $F(1,118)=5.33, p=0.023$; switch X relation type $F(1,118)=9.45, p=0.003$; target X relation type: $F(1,118)=7.84, p=0.006$; and switch x relation type x modality: $F(1,118)=4.79, p=0.031$. To interpret these interactions in more detail, RTs for relation type and modality variables are presented for all prime-target language combinations in Graph 1. RT values for morphologically related targets are presented in red, while RT values for unrelated pairs are presented in blue.

GRAPH 1 – Reaction times for variables relation type and modality per language combination: RTs for targets preceded by an auditory prime presented on the left (AUD) of each graph; RTs for targets preceded by written primes are presented on the right of each graph. RTs for morphological pairs are presented in blue; RTs for unrelated pairs are presented in red. Graphs a and b present RTs for the non-switch condition with L1 and L2 as target languages, respectively; Graphs c and d present RTs for the switch condition with L1 and L2 as target languages, respectively



In Graph 1a, we can see that there was a statistically significant morphological priming effect for the non-switch condition with L1 as the target language, i.e. L1 to L1 priming, and only when the prime was auditory. The priming effect was expressed by faster RTs for morphologically related targets as compared to unrelated targets: morph. 700.44ms(161.41ms) vs. unrel. 744.21ms(150.77ms), with p-value of $p=0.016$. For all other language combinations in different modality combinations there was no statistically significant effect (for all pairwise comparisons see Table 2).

For the switch direction L1 to L2 in the auditory condition, morphologically related targets yielded far slower RTs than for unrelated targets (see Graph 1b): morph. 856.49ms(204.24ms) vs. unrel. 766.60ms(176.50ms), $p<0.001$.

For most language combinations, targets preceded by auditory primes yielded slower RTs than targets preceded by written primes, especially when unrelated targets are compared, with statistically significant values for L1 to L1 ($p=0.003$), L2 to L2 ($p=0.048$), and, L2 to L1 (0.015). RTs for auditory modalities also yielded slower RTs as compared to written modalities for morphological pairs, but only in the switch conditions: L1 to L2 (AUD. 726.69ms(183.01ms) vs. WR. 672.61ms(156.54ms), $p=0.015$), and L2 to L1 (AUD. 726.69ms(183.01ms) vs. WR. 672.61ms(156.54ms), $p=0.340$). In the latter switch condition, the apparent difference was not significant statistically, probably due to relatively high standard error values as visualized by the errors bars in Graph 1d.

TABLE 2 – Pairwise comparisons for relation type and modality variables per language combination: Reaction times are given in ms with standard error between parentheses. P values were calculated for pairwise comparisons applying an LSD test

	L prime - L target	relation type	modality (RTs in ms)		<i>p value</i>
			AUD	WR	
non-switch	L1 to L1	morph.	700.44 (161.42)	669.01 (163.15)	<i>0.130</i>
		unrel.	744.12 (150.77)	694.19 (165.31)	<i>*0.003</i>
		<i>p values</i>	<i>*0.016</i>	<i>0.236</i>	
	L2 to L2	morph.	770.81 (161.46)	738.33 (171.52)	<i>0.096</i>
		unrel.	792.52 (177.02)	744.37 (180.41)	<i>*0.048</i>
		<i>p values</i>	<i>0.312</i>	<i>0.773</i>	
switch	L2 to L1	morph.	726.69 (183.01)	672.61 (156.54)	<i>*0.002</i>
		unrel.	732.14 (158.40)	690.79 (157.31)	<i>*0.015</i>
		<i>p values</i>	<i>0.766</i>	<i>0.351</i>	
	L1 to L2	morph.	856.48 (204.24)	784.29 (161.02)	<i>0.340</i>
		unrel.	766.60 (176.50)	769.51 (149.61)	<i>0.858</i>
		<i>p values</i>	<i>*<0.001</i>	<i>0.403</i>	

4.2 Behavioral Data: conclusion

The data show that this group of late bilinguals does not present any morphological priming effect in L2. It is difficult to say whether this effect can be explained by an overall slower processing in L2 or whether this is related specifically to difficulty processing morphologically complex words in L2. Both RTs for morphologically related and unrelated targets were equally slower in L2 as compared to L1. In L1 there was a facilitation effect due to morphological priming, but this was only

statistically significant for targets that followed auditory primes. This seems to suggest that morphological priming occurs irrespective of an orthographic repetition effect, but this conclusion is tentative given that morphological priming only robustly occurred with cross-modal priming, thus offering no grounds for comparison with unimodal priming. Across the board, cross-modal conditions resulted in higher RTs, indicative of a cognitive cost.

There was also no cross-linguistic morphological priming. On the contrary, the highest RTs overall were in response to morphologically related targets preceded by auditory pairs in the L2 to L1 condition. Switching from L1 to L2 yielded higher RTs across all conditions, but the interesting point here is that RTs were significantly higher for morphologically related targets than for unrelated targets in the same language switching condition. This seems in line with what the Revised Hierarchical Model (RHM) would predict; namely, that relatively rich semantic features activated by L1 primes co-activate many L1 related candidates, but not an equal amount of L2 candidates, hindering lexical access in L2. For unimodal written pairs in the L1 to L2 direction, the difference between RTs for morphologically related and unrelated targets was not statistically significant, but there is a tendency towards an asymmetric effect also.

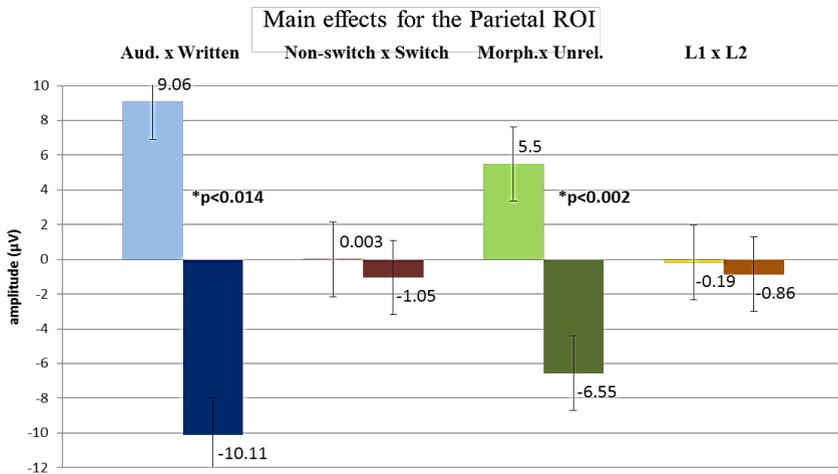
4.3 ERP data: results

In this analysis we will discuss mean amplitudes, measured in μ Volts, extracted from the ERP waveforms within the 350-450ms interval after word presentation onset (LUCK, 2005). An ANOVA analysis for repeated measures was carried out, considering the following variables: (i) switch: non switch vs. switch; (ii) relation type: morphological vs. unrelated; (iii) target language: L1 vs. L2; (iv) modality: auditory vs. written; and (v) ROI: Frontal, Central, Temporal, Parietal, and Occipital. The analysis revealed main effects for relation type: $F(1,13)=18.21$, $p=0.001$, modality: $F(1,13)=13.18$, $p=0.003$, and ROI($4,52$)= 8.16 , $p<0.001$. There were no main effects for the switch variable ($F(1,13)=0.014$, $p=0.91$) and target language ($F(1,13)=1.03$, $p=0.33$). There were no interactions between the variable ROI and any of the experimental variables (ROI vs. target L.: $F(4,52)=0.13$, $p=0.97$;

ROI vs. mod: $F(4,52)=0.53$, $p=0.72$; ROI vs. rel.type: $F(4,52)=1.44$, $p=0.23$; ROI vs. switch: $F(4,52)=0.75$, $p=0.56$). The fact that there was no interaction between ROI and any of the variables, may be taken as an indication that the main effect of ROI as a variable is mostly caused by differences in the strength of the responses in different ROIs, rather than by any experimental effect that would be specific for only one of the ROIs. Reports in the literature cite central and parietal areas as most consistent for the N400 effect (KUTAS; FERERMEIER, 2000; LUCK, 2005; LAU *et al.*, 2008). Therefore, both our analysis and findings in literature support the selection of the parietal ROI as representative for N400 effects in order to illustrate the results of our analysis.

An ANOVA analysis for repeated measures on mean amplitudes only from the Parietal ROI showed that main effects in this region mirror those described in the previous paragraph. In Graph 2, mean amplitude measures are shown for the 4 experimental variables. Negative amplitudes as compared to relatively less negative amplitudes or positive amplitudes are considered to reflect a higher cognitive cost. For example, for the modality variable there was a main effect ($F(1,14)=7.79$, $p=0.014$, which was driven by more negative amplitudes for targets preceded by written primes ($-10.11\mu\text{V}$) than for targets preceded by auditory primes ($9.06\mu\text{V}$). Thus, the data seem to suggest that the processing of written targets takes more cognitive effort relative to the processing of auditory targets. The main effect for relation type ($F(1,14)=14.84$, $p=0.002$) revealed that overall targets of unrelated pairs elicited more negative amplitudes ($-6.55\mu\text{V}$) than targets of morphologically related pairs ($5.5\mu\text{V}$). There were no effects for the target language variable ($F(1,14)=0.014$, $p=0.91$) nor for the switch variable ($F(1,14)=0.08$, $p=0.784$), indicating that overall there were no significant differences between targets in L1 and L2, irrespective of the previous language context (L1 or L2). The lack of interaction between these variables confirms this interpretation ($F(1,14)=0.78$, $p=0.39$).

GRAPH 2 – Main effects for the parietal ROI: mean amplitudes in μV are shown for modality in blue, switch in purple, relation type in green and target language in orange



To examine the effects of the experimental conditions on the waveforms, Figures 3a and b are presented with ERPs for target language L1 (Fig. 3a) and L2 (Fig. 3b), collapsing switch and non-switch conditions so as to make the comparison between the four waveforms representing modality and relation type easier. There are two major peaks in these ERP graphs: the first peak around 170ms-200ms, and a second around 400ms. Our analysis focuses on mean amplitudes extracted from the 350-450ms interval as our interest is to examine the N400 component. Visual inspection of the waveforms in this interval (as marked by the shaded area) seems to suggest that there were priming effects for both L1 and L2, given that unrelated targets yielded higher amplitudes (i.e. more negative values, with negative values plotted upwards) as compared to morphologically related targets. This seems to be the case for targets preceded by auditory primes (black: morph vs. blue:unrel.) as well as for targets preceded by written primes (red: morph. vs. green: unrel.). The more negative (or less positive) amplitudes for unrelated targets in comparison to morphologically related targets may indeed imply there was a relative ease in processing as a consequence of morphological priming. However, post-hoc analyses (LSD) indicated that there was only

a (nearly) statistically significant morphological priming effect for the cross-modal condition with L1 as the target language. For this condition, unrelated targets yielded higher (i.e. less positive) amplitudes as compared to morphologically related targets: $0.25(34.78)$ vs. $17.44\mu\text{V}(38.68 \mu\text{V})$, $p=0.05$. Perhaps qualitatively apparent morphological priming effects for other modalities with L2 as a target language are not robust enough to survive statistical analysis, which may be due to the fact that standard deviations are high, indicative of great variation in the data (for all values see Table 3).

FIGURE 3– ERPs for the Parietal ROI comparing Target Language – Waveforms for targets of morphologically related pairs are in black when preceded by an auditory prime, and in red when preceded by a written prime. Waveforms for targets of unrelated related pairs are in blue when preceded by an auditory prime, and in green when preceded by a written prime. The shaded area indicates the interval (350–450ms) from which mean amplitudes were extracted

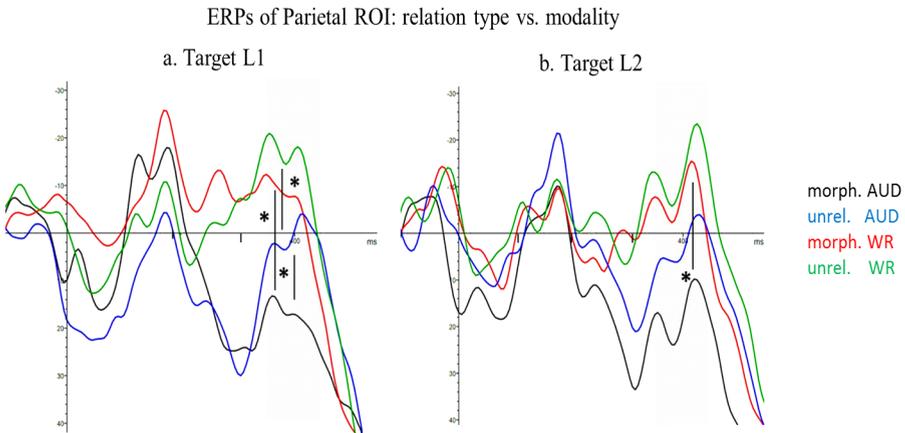


TABLE 3 – Post-hoc analysis of mean amplitudes for ERPs of the Parietal ROI: mean amplitudes indicated in μV with standard deviation in parentheses. Fisher's LSD (least significant difference) test was applied to carry out pairwise comparisons

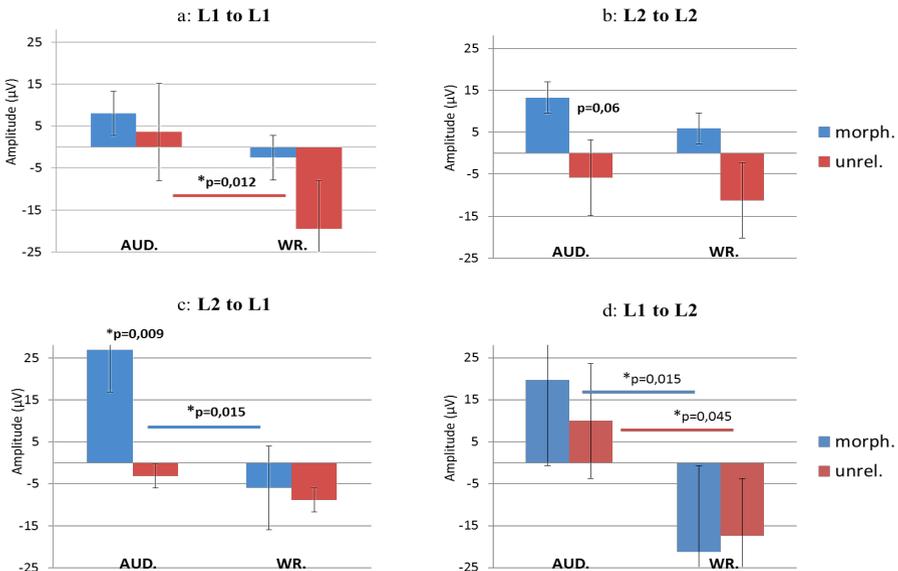
target language	relation type	modality (values in μV (std.dev))		post hoc comparison (LSD)
		AUD	WR	<i>p-value</i>
L1	morph.rel.	17.44(38.68)	-4.27(36.78)	0.02
	unrel.	0.25(34.78)	-14.18(42.07)	0.03
	<i>p-value</i>	0.05	0.24	
L2	morph.rel.	16.51(35.4)	-7.67(41.36)	0.02
	unrel.	2.05(45.45)	-14.31(41.22)	0.14
	<i>p-value</i>	0.12	0.37	

Tendencies for morphological priming could be observed irrespective of cross-modal priming. In fact, for cross-modal priming, all amplitudes were lower, and differences between unrelated and morphologically related targets were greater; so much so that only cross-modal priming led to a statistically more pronounced effect for L1 targets. To investigate further whether morphological priming carried over even in switching conditions, all language combinations were plotted in Graph 4. Pairwise comparisons revealed that the morphological priming effect presented in Figure 3 for L1 as the target language were most likely driven by positive amplitudes for morphologically related targets in the L2 to L1 switching condition, but only in the cross-modal condition (L2 to L1: morph. $26.87\mu\text{V}(8.57\mu\text{V})$ vs. unrel. $-3.09\mu\text{V}(10.25\mu\text{V})$, $p=0.009$). Morphological priming for L2 as the target language reached near significance with a p value of 0.06 for the L2 to L2 condition (morph. $13.23\mu\text{V}(9.58\mu\text{V})$ vs. unrel. $-5.88\mu\text{V}(11.35\mu\text{V})$). This confirms the interpretation that there was a tendency for morphological priming, but that it was not statistically robust. The data suggest that morphological priming can occur in L2, even in a cross-linguistic context. Moreover, this effect does not rely on an orthographic repetition effect given that the morphological priming effect was restricted to cross-modal morphological priming pairs.

For the L1 to L2 switching condition, there were no differences between the mean amplitudes of morphologically related and unrelated targets, but there was a statistically significant effect between unimodal and cross-modal priming, with more negative amplitudes for unimodal priming (morph.: AUD. 19.75 μ V(8.94 μ V) vs. WR. -21.27 μ V(10.95 μ V), $p=0.02$; and unrel.: AUD. 9.97 μ V(12.14 μ V) vs. WR. -17.35 μ V(8.80 μ V), $p=0.045$). Cross-modal priming also led to differences in other language combinations (L1L1 and unrel.: AUD. 3.59 μ V(7.78 μ V) vs. WR. -19.51 μ V(9.56 μ V), $p=0.012$; and L2L1 and morph.: AUD. 26.87 vs. WR: -6.0 μ V(10.6 μ V), $p=0.015$). Nonetheless, it was in the L1 to L2 switching conditions that these differences were most pronounced since they could be observed between both morphologically related and unrelated pairs.

4.3 ERP data: conclusions and comparison to RT data

GRAPH 4 – Mean amplitudes for variables relation type and modality per language combination - Mean amplitudes for targets preceded by an auditory prime presented on the left (AUD) of each graph; Mean amplitudes for targets preceded by written primes are presented on the right of each graph. Mean amplitude values for morphological pairs are presented in blue, for unrelated pairs in red. Graphs a and b present mean amplitudes for the non-switch condition with L1 and L2 as target languages, respectively; Graphs c and d present mean amplitudes for the switch condition with L1 and L2 as target languages, respectively



Lower amplitudes for morphologically related targets as compared to unrelated targets in the 350-450ms interval indicated that there was an overall morphological priming effect for targets in L1. The fact that this facilitation was largest for switching from L2 to L1 suggests that L2 can successfully map onto meaning, thus, easily activating L1 translation pairs that prime morphologically related words in L1. Nearly significant morphological priming effects for the non-switch condition in L2 (L2 to L2) seems to imply that late bilinguals can establish morphological relations between words in L2 and that morphological parsing, although not as fast and effective as in L1, can occur.

The fact that these morphological priming effects mostly occurred with cross-modal priming is somewhat surprising, given that priming effects tend to decrease in the context of cross-modal priming (HOLCOMB *et al.*, 2005; KIYONAGA *et al.*, 2007). However, data in the literature are based on monolingual processing. It might be that the conversion from orthographic forms to phonological representations becomes more costly when more than one language is involved. Each language has its own particular grapheme to phoneme mapping, some of which may even overlap or contrast between languages (for example, the grapheme 'g' maps onto the phoneme /g/ in PB, but onto the phoneme /χ/ in Dutch). Bearing that in mind, we may perhaps think of unimodal priming in language switch trials as something that can be quite costly. Perhaps the cognitive cost of this phenomenon even carried over to priming pairs that were monolingual since non-switching trials were alternated with switching trials in the same block. In any case, it seems that morphological priming effects in both L2 and L1 are not dependent on a mere orthographical repetition effect. Although without any robust priming effect in unimodal conditions for the purpose of comparison, this conclusion remains tentative.

Different from the ERP data, the behavioral data did not show any evidence of morphological priming effects in L2. This implies that RTs and ERP measures possibly reflect different stages of word recognition, and that they may be affected differently by task-related processing strategies. N400 components are known to be sensitive to lexical properties of roots (e.g. *√friend*) more so than to lexical properties of full word forms (e.g. *friendliness*) (PYLKKÄNEN *et al.*, 2000, 2002, 2003; FRANÇA *et al.*, 2008). Therefore, lower (or less negative) amplitudes in response to morphologically complex pairs in L2 (e.g.

amigo-amigável (“friend-friendly”)) may reflect morphological priming effects because bilinguals recognize the root embedded in the target word ($\sqrt{\text{friend}}$) for which the meaning was already activated from when the prime was processed. Thus facilitation is due to the sustained or repeated activation of the meaning representation of the root which is repeated for primes and targets in the same language. This means that, at least, morphologically complex words can be parsed for lexical roots and derivational suffixes in L2 processing. For late bilinguals, there is probably a qualitative difference in the semantic interpretation of roots and suffixes, where the semantic interpretation of lexical roots is sustained by word form to meaning mapping, while the interpretation of derivational suffixes depends on interpreting highly predictable and rule-based meaning added to a lexical root-based meaning (e.g. $\sqrt{\text{friend}} + \text{ly}$ (attribute) + *-ness* (quality/state of) \rightarrow *friendliness*: state of being friend-like). So-called superficial morphological processing in L2 is likely to affect subsequent processing of derivational suffixes more than lexical root to meaning mapping. This difficulty becomes apparent in higher RTs for morphologically related targets in L2, such that no facilitation effect can be observed in the RT analysis as opposed to analysis of the mean amplitudes, in which we do see a morphological priming effect. This account is in line with the Superficial Processing Hypothesis (CLAHSEN; FELSER, 2006; CLAHSEN *et al.*, 2010).

It seems less plausible to attribute this difficulty in morphological processing to an L1 bias for monomorphic processing given that there is a facilitation effect for morphological priming in L1, which is confirmed both by RT and mean amplitude measures, even though, in Dutch, morphologically complex words formed by derivation are less common than words formed by compounding.

The asymmetrical switch effect predicted by the Revised Hierarchical Model (RHM) was confirmed by the behavioral data. There was a significant difference between RT values for morphologically related targets and unrelated targets for cross-modal pairs in the L1 to L2 condition. RTs for morphologically related targets were significantly higher than those for unrelated pairs. This implies that especially related meanings caused a processing cost as opposed to unrelated meanings, which is something that is well accounted for by the RHM. This model attributes a special status to the relative fragility of concept to word form links in L2 as compared to word form to concept mapping, which is a

notion that explains the observed effects in the behavioral data well. For the ERP data, we did not observe a similar asymmetrical priming effect for the L1 to L2 condition. This can be explained if we consider once again that N400 components reflect the activation of lexical root based meaning, and that there is a difference in L2 between the semantic interpretation of roots, which is relatively easy compared to interpreting derivationally generated semantics in L2. It is exactly this grammatically constrained rule based processing that is difficult to learn for late bilinguals due to cognitive-biological limitations for language acquisition beyond the critical period.

Bibliographic references

BHATIA, T. K.; WILLIAM, C. R. Introduction. In: BHATIA, T. K.; WILLIAM, C. R. (Ed.). *The Handbook of Bilingualism*. Malden, MA: Blackwell Publishing, 2005.

BORODITSKY, L. Linguistic Relativity. In: NADEL, L. (Ed.). *Encyclopedia of Cognitive Science*. London, UK: MacMillan Press, 2003. p. 917-921.

BRYSBART, M.; DIJKSTRA, T. Changing views on word recognition in bilinguals. In: MORAIS, J.; d'YDEWALLE, G. (Ed.). *Bilingualism and second language acquisition*. Belgium: Royal Academes for Science and the Arts of Belgium, 2006.

BULTENA, S. Bilingual visual word recognition and lexical access. In: KROLL, J. F.; De GROOT, A. (Ed.). *Handbook of bilingualism: psycholinguistic approaches*. New York: Oxford University Press, 2005. p. 179-201.

CLAHSEN, H.; FELSER, C. Grammatical processing in language learners. *Applied Psycholinguistics*, Cambridge University Press, v. 27, p. 3-42, 2006.

CLAHSEN, H.; FELSER, C.; NEUBAUER, K.; SATO, M.; SILVA, R. Morphological structure in Native and Non-Native Language Processing. *Language Learning*, Wiley Online Library, v. 60, p. 21-43, 2010.

DE GROOT, A. M. B. La représentation lexico-sémantique et l'accès lexical chez le bilingue. *Psychologie Française*, Elsevier, v. 43, p. 297-312, 1998.

DE GROOT, A. M. B.; DELMAAR, P.; LUPKER, S. J. The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology*, Taylor & Francis Online, v. 53, n. 2, p. 397-428-432, 2000.

DÍAZ, B.; ERDOCIA, K.; de MENEZES, R. F.; MUELLER, J. L.; SEBASTIÁN-GALLÉS, N.; LAKA, I. Electrophysiological correlates of second-language syntactic processes are related to native and second language distance regardless of age of acquisition. *Frontiers in Psychology*, Frontiers Media, v. 7, n.133, 2016. <https://doi.org/10.3389/fpsyg.2016.00133>

DIEPENDAELE, K.; DUÑABEITUA, J. A.; MORRIS, J.; KEULEERS, E. Fast morphological effects in first and second language word recognition. *Journal of Memory and Language*, Elsevier, v. 64, n. 4, p. 344-358, 2011. <https://doi.org/10.1016/j.jml.2011.01.003>

DIJKSTRA, T.; HEUVEN, W. J. B. van. The BIA model and bilingual word recognition In: J. GRAINGER, J.; JACOBS, A. M. (Ed.). *Localist connectionist approaches to human cognition*. Mahwah, N.J.: Erlbaum, 1998. p. 189-225.

DIJKSTRA, T.; GRAINGER, J.; HEUVEN, W. J. B. van. Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, Elsevier, v. 41, Issue 4, p. 496-518, 1999. <https://doi.org/10.1006/jmla.1999.2654>

DIJKSTRA, T.; MOSCOSO DEL PRADO MARTIN, F.; SCHULPEN, P.; SCHREUDER, R.; BAAYEN, R.H. A roommate in cream: morphological family size effects on interlingual homograph recognition. *Language & Cognition Process*, Taylor & Francis Online, v. 20, n. 1/2, p. 7-41, 2005. <https://doi.org/10.1080/01690960444000124>

DOM, A.; VEJA, M. De; BARBER, H. Event-related Brain Potentials Elicited by Morphological, Homographic, Orthographic, and Semantic Priming. *Journal of Cognitive Neuroscience*, MIT Press, v. 16, n. 4, 598-608, 2004. <https://doi.org/10.1162/089892904323057326>

DOMINGUEZ, A.; SEGUÍ, J.; QUETOS, F. The time course of inflectional morphological priming. *Linguistics*, De Gruyter, v. 40, p. 235-259, 2002.

FELDMAN, L. B. Beyond orthography and phonology: differences between inflections and derivations. *Journal of Memory and Language*, Elsevier, v. 33, p. 442-470, 1994.

FELDMAN, L. B.; KOSTIC, A.; BASNIGHT-BROWN, D. M.; FILIPOVIC-DURDEVIC, D.; PASTIZZO, M. J. Morphological facilitation for regular and irregular verb formations in native and non-native speakers: little evidence for two distinct mechanisms. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 13, p. 119-135, 2009.

FRANÇA, I.; LEMLE, M.; GESUALDI, A. M.; CAGY, M.; INFANTOSI, A. F. C. A neurofisiologia do acesso lexical: palavras em português. *Veredas*, Programa de Pós-Graduação em Linguística da UFJF, v. 12, n. 2, p. 34-49, 2008.

GARCIA, D. C. de; MAIA, M. A. R.; FRANÇA, A. I. The time course of word recognition: evidence from Brazilian Portuguese, *ReVEL - Revista Virtual de Estudos da Linguagem*, [s.l.], v. 10, p. 169-181, 2012.

GREEN, D. W. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 1, n. 1, p. 67-81, 1998. DOI: <https://doi.org/10.1017/S1366728998000133>.

GROSJEAN, F. Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36, 3-15, 1989. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5)

GROSJEAN, F. Studying bilinguals: Methodological and conceptual issues. *Bilingualism, Language and Cognition*, Cambridge University Press, v. 1, n. 2, p. 131-149, 1998. DOI: <https://doi.org/10.1017/S136672899800025X>.

GROSJEAN, F. Individual bilingualism. In: ASHER, R. E. (Ed.). *The encyclopaedia of language and linguistics*. Oxford, UK: Pergamon Press, 1994. p. 1656-1660.

HAHNE, A.; FRIEDERICI, A. D. Processing a second language: late learners' comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 4, Issue 2, p 123-141, 2001.

HAHNE, A.; MUELLER, J. H.; CLAHSEN, H. Morphological processing in a second language: Behavioural and event-related potential evidence for storage and decomposition, *Journal of Cognitive Neuroscience*, MIT Press, v. 18, p. 121-134, 2006. <https://doi.org/10.1162/089892906775250067>

HALLE, M.; MARANTZ, A. Distributed Morphology and the pieces of inflection. In: HALE, K.; KEYSER, S. J. (Ed.). *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: The MIT Press, 1993. p. 111-176.

HEREDIA, R. R.; BROWN, J. M. Bilingual memory. In: BHATIA, T. K.; RITCHIE, W. C.; (Ed.). *Handbook of Bilingualism*. Malden, MA: Blackwell Publishers, 2004.

HOLCOMB, P. J.; ANDERSON, J.; GRAINGER, J. An electrophysiological study of cross-modal repetition priming. *Psychophysiology*, Wiley Online Library, v. 42, n. 5, p. 493-507, 2005. <https://doi.org/10.1111/j.1469-8986.2005.00348.x>

JACQUET, M.; FRENCH, R. M. The BIA++: Extending the BIA+ to a dynamical distributed connectionist framework. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 5, n. 3, p. 202-205, 2002. <https://doi.org/10.1017/S1366728902223019>

KIYONAGA, K.; GRAINGER, J.; MIDGLEY, K.; HOLCOMB, P. J. Masked cross-modal repetition priming: an event-related potential investigation. *Language and Cognitive Processes*, Taylor & Francis Online, v. 22, n. 3, p. 337-376, 2007. <https://doi.org/10.1080/01690960600652471>

KOTZ, S. A. Neurolinguistic evidence for bilingual language representation: a comparison of reaction times and event-related brain potentials. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 4, n. 2, p. 143-154, 2001. <https://doi.org/10.1017/S1366728901000244>

KOTZ, S. A.; ELSTON-GÜTTLER, K. The role of proficiency in processing categorical and associative information in the L2 as revealed by reaction times and event-related brain potentials. *Journal of Neurolinguistics*, Elsevier, v. 17, Issues 2/3, p. 215-235, 2004.

KOTZ, S. A.; HOLCOMB, P. J.; OSTERHOUT, L. ERPs reveal comparable syntactic sentence processing in native and non-native readers of English. *Acta Psychologica*, Elsevier, v. 128, n. 3, p. 514-527, 2008. <https://doi.org/10.1016/j.actpsy.2007.10.003>

KROLL, J. F.; de GROOT, A. M. B. Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In: de GROOT, M. B.; KROLL, J. F. (Ed.). *Tutorials in bilingualism*. Mahwah, NJ: Erlbaum, 1997. p 169-199.

KROLL, J. F.; STEWART, E. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, Elsevier, v. 33, p. 149-174, 1994. <https://doi.org/10.1006/jmla.1994.1008>

KROLL, J. F.; HELL, J. G. van; TOKOWICZ, N.; GREEN, D. The revised hierarchical model: a critical review and assessment. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 13, n. 3, p. 373-381, 2010. <https://doi.org/10.1017/S136672891000009X>

KUTAS, M.; FEDERMEIER, K. D. Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, Elsevier, v. 4, n. 12, p. 463-470, 2000.

LAU, E.; PHILLIPS, C.; POEPPPEL, D. A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, Macmillan Publishers, v. 9, n. 12, p. 920-933, 2008. <https://doi.org/10.1038/nrn2532>

LEMHÖFER, K.; DIJKSTRA, T.; SCHRIEFERS, H.; BAAYEN, H. R.; GRAINGER, J.; ZWITSERLOOD, P. Native Language Influences on Word Recognition in a Second language: A Mega-Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, PubliMed, v. 34, p. 12-31, 2008.

LENSINK, S. E.; VERDONSCHOT, R. G.; SCHILLER, N. O. Morphological priming during language switching: an ERP study. *Frontiers in Human Neuroscience*, Frontiers Media, v. 8, p. 995, 2014. doi: 10.3389/fnhum.2014.00995. <https://doi.org/10.3389/fnhum.2014.00995>

LEWY, N.; GROSJEAN, F. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, Elsevier, v. 33, p. 149-174, 1994. <https://doi.org/10.1006/jmla.1994.1008>

LIPSKY, J. M. Code switching and bilingual competence. In: PARADIS M. (Ed.). *Aspects of bilingualism*. Columbia: Hornbeam Press, 1978. p. 250-264.

LUCK, S. *Introduction to the Event-Related Potential Technique*. Cambridge: MIT, 2005.

MARANTZ, A. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In: ANNUAL PENN LINGUISTICS COLLOQUIUM, 21st., 1997, Pennsylvania. *University of Pennsylvania Working Papers in Linguistics* (PWPL), Pennsylvania, v. 4, n. 2, p. 201-225, 1997.

MARANTZ, A. Generative linguistics within the cognitive neuroscience of language. *The Linguistic Review*, De Gruyter, v. 22, p. 429-445, 2005.

MEISEL, J. M. Getting FAT: Finiteness, agreement and tense in early grammars In: MEISEL, J. M. (Ed.). *Bilingual first language acquisition: French and German grammatical development*. Amsterdam: Benjamins, 1994. p. 89-130. <https://doi.org/10.1075/lald.7>

MUNTE, T. F.; SAY, T.; CLAHSEN, H.; SCHILTZ, K.; KUTAS, M. Decomposition of morphologically complex words in English: evidence from event-related brain potentials. *Brain Research. Cognitive Brain Research Journal*, PubMed, v. 7, n. 3, p. p. 241-253, Jan. 1999.

NIEDERMEYER, E.; SILVA, F. L. *Electroencephalography: Basic principles. Clinical applications and related fields*. 4. ed. Baltimore. Maryland: Williams & Wilkins, 1982. p. 1258.

OSTERHOUT, L.; MCLAUGHLIN, J.; BERSICK, M. Event-related brain potentials and human language. *Trends in Cognitive Sciences*, New York, v. 1, n. 6, p. 203-209, Sept. 1997. [https://doi.org/10.1016/S1364-6613\(97\)01073-5](https://doi.org/10.1016/S1364-6613(97)01073-5)

PARADIS, M. The cognitive neuropsychology of bilingualism. In: de GROOT, A.; KROLL, J. (Ed.). *Tutorials in bilingualism: psycholinguistic perspectives*. Hillsdale, NJ.: Lawrence Erlbaum, 1997. p. 331-354.

PORTIN, M.; LEHTONEN, M.; LAINE, M. Processing of inflected nouns in late bilinguals. *Applied Psycholinguistics*, Oxford Academic, v. 28, p. 135-156, 2007.

PYLKKÄNEN, L.; STRINGFELLOW, A.; FLAGG, A.; MARANTZ, A. Neural response sensitive to repetition and phonotactic probability: MEG investigations of lexical access. In: BIOMAG, 2000, Espoo, Finland. *Proceedings...* Espoo: Helsinki University of Technology, 2000. p. 363-367.

PYLKKÄNEN, L.; GONNERMAN, L.; STRINGFELLOW, A.; MARANTZ, A. *Disambiguating the source of phonological inhibition effects in lexical decision: an MEG study*. New York: New York University; Lehigh University; University of California at San Diego, 2003

PYLKKÄNEN, L.; STRINGFELLOW, A.; MARANTZ, A. Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language*, Elsevier, v. 81, n. 1/3, p. 666-678, 2002. <https://doi.org/10.1006/brln.2001.2555>

PYLKKÄNEN, L.; MARANTZ, A. Tracking the time course of word recognition with MEG, *Trends in Cognitive Sciences*, New York, v. 7, n. 5, p. 187-189, 2003. [https://doi.org/10.1016/S1364-6613\(03\)00092-5](https://doi.org/10.1016/S1364-6613(03)00092-5)

SCHRIEFERS, H.; FRIEDERICI, A.; GRAETZ, P. Inflectional and derivational morphology in the mental lexicon: Symmetries and asymmetries in repetition priming. *Quarterly Journal of Experimental Psychology*, ResearchGate, v. 44A, p. 373-390, 1992.

SILVA, R.; CLAHSEN, H. Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 11, p. 245-260, 2008.

SOTO, M. Interceptando mecanismos de alternância bilíngue: a micromodularidade revelada nos ERPs. Dissertação de mestrado, f. 185, UFRJ, Faculdade de Letras, Rio de Janeiro, 2010.

STOCKALL, L.; MARANTZ, A. A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, John Benjamins, v. 1, n. 1, p. 85-123, 2006. <https://doi.org/10.1075/ml.1.1.07sto>

TANNER, D.; MCLAUGHLIN, J.; HERSCHENSOHN, J.; OSTERHOUT, L. Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, Cambridge University Press, v. 16, p. 367-382, 2016. <https://doi.org/10.1017/S1366728912000302>

ULLMAN, M. Contributions of memory circuits to language: the declarative/procedural model. *Cognition*, Elsevier, v. 92, p. 231-270, 2004. <https://doi.org/10.1016/j.cognition.2003.10.008>

ULLMAN, M. A Cognitive neuroscience perspective on second language acquisition: the declarative/procedural model. In: SANZ, C. (Ed.). *Mind and context in adult second language acquisition: methods, theory and practice*. Washington D.C.: Georgetown University Press, 2005. p. 141-178.

WEBER-FOX, C. M.; NEVILLE, H. Maturation Constraints on Functional Specializations for Language Processing: ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, MIT Press, v. 8, n. 3, p. 231-256, 1996. <https://doi.org/10.1162/jocn.1996.8.3.231>

WHORF, B. (Ed.). *Language, thought, and reality: selected writings of Benjamin Lee Whorf*. New York; London: MIT Press, 1956.

APPENDIX 1 – Stimuli list

Series	L combination	relation type	modality	prime	(gram. class)	target	(gram. class)
1	L1-L1	morph.	AUD	tijd	NOUN	TIJDIG	ADJ
1	L1-L1	morph.	AUD	zorg	NOUN	ZORGELOOS	ADJ
1	L1-L1	morph.	AUD	moeder	NOUN	MOEDERLIJK	ADJ
1	L1-L1	morph.	AUD	mens	ADJ	MENSHEID	NOUN
1	L1-L1	morph.	AUD	kracht	NOUN	KRACHTIG	ADJ
1	L1-L1	morph.	AUD	voeden	VERB	VOEDZAAM	ADJ
1	L1-L1	morph.	AUD	zuiver	ADJ	ZUIVERHEID	NOUN
1	L1-L1	morph.	AUD	macht	NOUN	MACHTIG	ADJ
1	L1-L1	unrel.	AUD	zoeken	VERB	KINDEREN	NOUN
1	L1-L1	unrel.	AUD	vraag	NOUN	ARTIKEL	NOUN
1	L1-L1	unrel.	AUD	slim	NOUN	LANGZAAM	ADJ
1	L1-L1	unrel.	AUD	kerk	NOUN	PLEZIER	NOUN
1	L1-L1	unrel.	AUD	middag	NOUN	RELATIE	NOUN
1	L1-L1	unrel.	AUD	knuffel	NOUN	KLACHTEN	NOUN
1	L1-L1	unrel.	AUD	klap	NOUN	BALKON	NOUN
1	L1-L1	unrel.	AUD	arbeid	NOUN	SPINAZIE	NOUN
1	L1-Hung		AUD	licht	ADJ	AGGÓDJÁL	
1	L1-Hung		AUD	graag	ADJ	ALAPZAT	
1	L1-Hung		AUD	dorst	NOUN	BEKÖTŐ	
1	L1-Hung		AUD	beurs	NOUN	BESÚGÓ	
1	L1-Hung		AUD	vloer	NOUN	BÓKJÁT	
1	L1-Hung		AUD	slank	ADJ	BOLOND	
1	L1-Hung		AUD	groente	NOUN	CSAHOL	
1	L1-Hung		AUD	ketting	NOUN	CSALÁS	
1	L1-Hung		AUD	sport	NOUN	CSAPÁS	
1	L1-Hung		AUD	functie	NOUN	CSOPORT	
1	L1-Hung		AUD	metode	NOUN	CSORDA	
1	L1-Hung		AUD	kozijn	NOUN	CSUPÁN	
1	L1-Hung		AUD	knoop	NOUN	DURRAN	
1	L1-Hung		AUD	plank	NOUN	EGYIK	

1	L1-Hung		AUD	braaf	ADJ	ELHELYEZ	
1	L1-Hung		AUD	nagel	NOUN	ELILLAN	
2	L2-L1	morph.	AUD	jogo	NOUN	SPELER	NOUN
2	L2-L1	morph.	AUD	seguir	VERB	VOLGZAAM	ADJ
2	L2-L1	morph.	AUD	brilho	ADJ	GLANSLOOS	ADJ
2	L2-L1	morph.	AUD	casar	VERB	TROUWERIJ	NOUN
2	L2-L1	morph.	AUD	sentir	VERB	VOELBAAR	ADJ
2	L2-L1	morph.	AUD	fome	NOUN	HONGERIG	ADJ
2	L2-L1	morph.	AUD	chorar	VERB	HUILERIG	ADJ
2	L2-L1	morph.	AUD	culpa	NOUN	SCHULDIG	ADJ
2	L2-L1	unrel.	AUD	comprar	VERB	GESPREK	NOUN
2	L2-L1	unrel.	AUD	conselho	NOUN	BLIJVEN	VERB
2	L2-L1	unrel.	AUD	andar	VERB	STRATEN	NOUN
2	L2-L1	unrel.	AUD	morte	NOUN	STRAND	NOUN
2	L2-L1	unrel.	AUD	rotina	NOUN	VRIENDIN	NOUN
2	L2-L1	unrel.	AUD	esconder	VERB	OCHTEND	NOUN
2	L2-L1	unrel.	AUD	pato	NOUN	STERKTE	NOUN
2	L2-L1	unrel.	AUD	trato	NOUN	HORLOGE	NOUN
2	L2-Hung		AUD	quieto	ADJ	ELÉRTE	
2	L2-Hung		AUD	lima	NOUN	HOZZÁJUK	
2	L2-Hung		AUD	fechar	VERB	LEGALÁBB	
2	L2-Hung		AUD	mestre	NOUN	IDEGESÍT	
2	L2-Hung		AUD	virar	VERB	HIÁNYA	
2	L2-Hung		AUD	colo	NOUN	HIRTELEN	
2	L2-Hung		AUD	intenso	ADJ	MONDUNK	
2	L2-Hung		AUD	urso	NOUN	ELEGET	
2	L2-Hung		AUD	anular	VERB	ELTÖRPÜL	
2	L2-Hung		AUD	maçã	NOUN	HISZEN	
2	L2-Hung		AUD	teimoso	ADJ	MAGYAR	
2	L2-Hung		AUD	morder	VERB	GYŰJTI	
2	L2-Hung		AUD	pulmão	NOUN	ÁTNÉZTE	
2	L2-Hung		AUD	furo	NOUN	OLDALON	
2	L2-Hung		AUD	umido	ADJ	MOZGÁST	

2	L2-Hung		AUD	labio	NOUN	EZENTÚL	
3	L1-L2	morph.	AUD	groot	ADJ	GRANDEZA	NOUN
3	L1-L2	morph.	AUD	kleur	NOUN	CORANTE	ADJ
3	L1-L2	morph.	AUD	begin	NOUN	INICIANTE	ADJ
3	L1-L2	morph.	AUD	klaar	ADJ	PRONTIDÃO	NOUN
3	L1-L2	morph.	AUD	zacht	ADJ	SUAVEMENTE	NOUN
3	L1-L2	morph.	AUD	droog	ADJ	SECADOR	NOUN
3	L1-L2	morph.	AUD	woede	NOUN	RAIVOSO	ADJ
3	L1-L2	morph.	AUD	zonde	NOUN	PECADO	NOUN
3	L1-L2	unrel.	AUD	alleen	ADJ	POSSÍVEL	ADJ
3	L1-L2	unrel.	AUD	niets	NOUN	DIFERENTE	ADJ
3	L1-L2	unrel.	AUD	park	NOUN	BARATO	ADJ
3	L1-L2	unrel.	AUD	scherm	NOUN	PRÉDIO	NOUN
3	L1-L2	unrel.	AUD	wetten	NOUN	VESTIDO	NOUN
3	L1-L2	unrel.	AUD	rivier	NOUN	CURIOSO	ADJ
3	L1-L2	unrel.	AUD	rust	NOUN	GELADO	ADJ
3	L1-L2	unrel.	AUD	winnen	VERB	GRÁVIDA	ADJ
3	L1-Hung		AUD	eigen	ADJ	ESÓÁLLÓ	
3	L1-Hung		AUD	kies	VERB	FEJEZET	
3	L1-Hung		AUD	keuken	NOUN	FEKHELY	
3	L1-Hung		AUD	geest	NOUN	FÉLCIPŐ	
3	L1-Hung		AUD	nuttig	ADJ	FOGDOS	
3	L1-Hung		AUD	beker	NOUN	FÖLDÚT	
3	L1-Hung		AUD	schrift	NOUN	FÜGGÖNY	
3	L1-Hung		AUD	blond	ADJ	FÚJTAK	
3	L1-Hung		AUD	type	NOUN	FÜLLEDT	
3	L1-Hung		AUD	avontuur	NOUN	GÉPÁGY	
3	L1-Hung		AUD	belg	NOUN	GYÁVÁVÁ	
3	L1-Hung		AUD	vest	NOUN	GYENGE	
3	L1-Hung		AUD	vlinder	NOUN	GYERMEK	
3	L1-Hung		AUD	erven	VERB	GYÖTÖR	
3	L1-Hung		AUD	pols	NOUN	HADIÜZEM	
3	L1-Hung		AUD	moraal	NOUN	HAJÓJÁT	

4	L2-L2	morph.	AUD	conta	NOUN	CONTADOR	NOUN
4	L2-L2	morph.	AUD	alegre	ADJ	ALEGRIA	NOUN
4	L2-L2	morph.	AUD	dúvida	NOUN	DUVIDOSO	ADJ
4	L2-L2	morph.	AUD	cabeça	NOUN	CABEÇADA	NOUN
4	L2-L2	morph.	AUD	parar	VERB	PARADA	NOUN
4	L2-L2	morph.	AUD	correr	VERB	CORRERIA	NOUN
4	L2-L2	morph.	AUD	ácido	ADJ	ACIDEZ	NOUN
4	L2-L2	morph.	AUD	cansar	VERB	CANSADO	ADJ
4	L2-L2	unrel.	AUD	grupo	NOUN	JAQUETA	NOUN
4	L2-L2	unrel.	AUD	letra	NOUN	MOCHILA	NOUN
4	L2-L2	unrel.	AUD	amanhã	NOUN	ESCUTAR	VERB
4	L2-L2	unrel.	AUD	selo	NOUN	FELIZES	ADJ
4	L2-L2	unrel.	AUD	raiz	NOUN	BONITO	ADJ
4	L2-L2	unrel.	AUD	queijo	NOUN	PALAVRA	NOUN
4	L2-L2	unrel.	AUD	burro	ADJ	SEGURO	ADJ
4	L2-L2	unrel.	AUD	joelho	NOUN	ESTRANHO	ADJ
4	L2-Hung		AUD	idéia	NOUN	LESÜTÖTT	
4	L2-Hung		AUD	topo	NOUN	LESZID	
4	L2-Hung		AUD	mesa	NOUN	LEVÉLSZÓ	
4	L2-Hung		AUD	natural	ADJ	LEZUHAN	
4	L2-Hung		AUD	triste	ADJ	LÖKÖTT	
4	L2-Hung		AUD	praga	NOUN	LOVASSÁG	
4	L2-Hung		AUD	dama	NOUN	MACSKÁM	
4	L2-Hung		AUD	leñçol	NOUN	MEGCSAL	
4	L2-Hung		AUD	álcool	NOUN	IKONRA	
4	L2-Hung		AUD	pera	NOUN	KITÉRTEK	
4	L2-Hung		AUD	ralo	NOUN	EGYMÁS	
4	L2-Hung		AUD	rastro	NOUN	EGYBEN	
4	L2-Hung		AUD	calar	VERB	KUTATÁST	
4	L2-Hung		AUD	mover	VERB	SZÁJÁBA	
4	L2-Hung		AUD	álgebra	NOUN	OKOKBÓL	
4	L2-Hung		AUD	irritar	VERB	KICSIT	
5	L1-L1	morph.	WR	land	NOUN	LANDELIJK	ADJ

5	L1-L1	morph.	WR	werken	VERB	WERKELOOS	ADJ
5	L1-L1	morph.	WR	wassen	VERB	WASSERIJ	NOUN
5	L1-L1	morph.	WR	slapen	VERB	SLAPERIG	ADJ
5	L1-L1	morph.	WR	pijn	NOUN	PIJNLIJK	ADJ
5	L1-L1	morph.	WR	zicht	NOUN	ZICHTBAAR	ADJ
5	L1-L1	morph.	WR	zuiden	NOUN	ZUIDELIJK	ADJ
5	L1-L1	morph.	WR	vrede	NOUN	VREDIG	ADJ
5	L1-L1	unrel.	WR	nieuw	ADJ	RECHTER	NOUN
5	L1-L1	unrel.	WR	fietsen	VERB	VERHAAL	NOUN
5	L1-L1	unrel.	WR	klein	ADJ	ZILVER	NOUN
5	L1-L1	unrel.	WR	lenen	VERB	DOCHTER	NOUN
5	L1-L1	unrel.	WR	komen	VERB	WONDER	NOUN
5	L1-L1	unrel.	WR	geheim	NOUN	WERELD	NOUN
5	L1-L1	unrel.	WR	katoen	NOUN	BELOFTE	NOUN
5	L1-L1	unrel.	WR	eten	VERB	HERFST	NOUN
5	L1-Hung		WR	plan	NOUN	HATÁROS	
5	L1-Hung		WR	donker	ADJ	HÉTALVÓ	
5	L1-Hung		WR	donker	ADJ	HÉTALVÓ	
5	L1-Hung		WR	luxe	NOUN	HITELBE	
5	L1-Hung		WR	avond	NOUN	HOZTUNK	
5	L1-Hung		WR	dalen	VERB	HULLADÉK	
5	L1-Hung		WR	focus	NOUN	IDŐJÁRÁS	
5	L1-Hung		WR	slecht	ADJ	IGAZÁN	
5	L1-Hung		WR	tempel	NOUN	IGAZGATÓ	
5	L1-Hung		WR	kist	NOUN	ILLETVE	
5	L1-Hung		WR	balen	VERB	IMÁDNI	
5	L1-Hung		WR	spin	NOUN	IRÁNYÁT	
5	L1-Hung		WR	touw	NOUN	ISMERT	
5	L1-Hung		WR	slap	ADJ	ISZÁKOS	
5	L1-Hung		WR	klep	NOUN	JÁTSZIK	
5	L1-Hung		WR	zweten	VERB	JÉGNEK	
6	L2-L1	morph.	WR	fala	NOUN	SPRAAKZAAM	
6	L2-L1	morph.	WR	viver	VERB	LEVENDIG	

6	L2-L1	morph.	WR	sonho	NOUN	DROMERIG	
6	L2-L1	morph.	WR	perigo	NOUN	GEVAARLIJK	
6	L2-L1	morph.	WR	sabor	NOUN	SMAKELIJK	
6	L2-L1	morph.	WR	dano	VERB	SCHADELIJK	
6	L2-L1	morph.	WR	luta	NOUN	VECHTER	
6	L2-L1	morph.	WR	quebrar	VERB	BREEKBAAR	
6	L2-L1	unrel.	WR	semana	NOUN	LASTIG	
6	L2-L1	unrel.	WR	cara	NOUN	DENKEN	
6	L2-L1	unrel.	WR	sentido	NOUN	WINTER	
6	L2-L1	unrel.	WR	vigor	NOUN	STOELEN	
6	L2-L1	unrel.	WR	pele	NOUN	BENIEUWD	
6	L2-L1	unrel.	WR	lápis	NOUN	GRAPPIG	
6	L2-L1	unrel.	WR	falha	NOUN	TEVREDEN	
6	L2-L1	unrel.	WR	honrado	ADJ	LAARZEN	
6	L2-Hung		WR	centro	NOUN	MEGLÉP	
6	L2-Hung		WR	arte	NOUN	MEGMENTI	
6	L2-Hung		WR	valer	ADJ	MEGŐRÜL	
6	L2-Hung		WR	simples	ADJ	MEGVESZ	
6	L2-Hung		WR	ativo	ADJ	MIKŐZBEN	
6	L2-Hung		WR	raça	NOUN	NEKIRONT	
6	L2-Hung		WR	pano	NOUN	NINCSEN	
6	L2-Hung		WR	lago	NOUN	NÖVÉNY	
6	L2-Hung		WR	calma	ADJ	NYAGGAT	
6	L2-Hung		WR	maré	NOUN	OLDALAK	
6	L2-Hung		WR	brisa	NOUN	PÉLDÁUL	
6	L2-Hung		WR	pintor	NOUN	ROMJAI	
6	L2-Hung		WR	honesto	ADJ	SZAVAIT	
6	L2-Hung		WR	garra	NOUN	TOVÁBBI	
6	L2-Hung		WR	cabra	NOUN	LÁTVÁNY	
6	L2-Hung		WR	cenoura	NOUN	LEÁNYKA	
7	L1-L2	morph.	WR	openen	VERB	ABERTURA	
7	L1-L2	morph.	WR	snel	ADJ	RAPIDEZ	
7	L1-L2	morph.	WR	vrij	ADJ	LIVREMENTE	

7	L1-L2	morph.	WR	lelijk	ADJ	FEIOSO	
7	L1-L2	morph.	WR	einde	NOUN	FINALMENTE	
7	L1-L2	morph.	WR	lachen	NOUN	RISADA	
7	L1-L2	morph.	WR	meten	VERB	MEDIÇÃO	
7	L1-L2	morph.	WR	traag	ADJ	LENTIDÃO	
7	L1-L2	unrel.	WR	zitten	VERB	LARANJA	
7	L1-L2	unrel.	WR	groen	ADJ	FÉRIAS	
7	L1-L2	unrel.	WR	stappen	VERB	SURPRESA	
7	L1-L2	unrel.	WR	gang	NOUN	LEVANTAR	
7	L1-L2	unrel.	WR	leider	NOUN	ESCOVA	
7	L1-L2	unrel.	WR	kast	NOUN	CORUJA	
7	L1-L2	unrel.	WR	staal	NOUN	OMBROS	
7	L1-L2	unrel.	WR	doel	NOUN	CARENTE	
7	L1-Hung		WR	muziek	NOUN		
7	L1-Hung		WR	lijst	NOUN		
7	L1-Hung		WR	thuis	NOUN		
7	L1-Hung		WR	klant	NOUN		
7	L1-Hung		WR	export	NOUN		
7	L1-Hung		WR	boete	NOUN		
7	L1-Hung		WR	acteur	NOUN		
7	L1-Hung		WR	grens	NOUN		
7	L1-Hung		WR	tegels	NOUN		
7	L1-Hung		WR	schaap	NOUN		
7	L1-Hung		WR	drop	NOUN		
7	L1-Hung		WR	koorts	NOUN		
7	L1-Hung		WR	happen	VERB		
7	L1-Hung		WR	spek	NOUN		
7	L1-Hung		WR	rommel	NOUN		
7	L1-Hung		WR	leugen	NOUN		
8	L2-L2	morph.	WR	fraco	ADJ	FRAQUEZA	NOUN
8	L2-L2	morph.	WR	corpo	NOUN	CORPORAL	ADJ
8	L2-L2	morph.	WR	peso	NOUN	PESADO	ADJ
8	L2-L2	morph.	WR	doce	ADJ	DOÇURA	NOUN

8	L2-L2	morph.	WR	espaço	NOUN	ESPACIAL	ADJ
8	L2-L2	morph.	WR	louco	ADJ	LOUCURA	NOUN
8	L2-L2	morph.	WR	coragem	NOUN	CORAJOSO	ADJ
8	L2-L2	morph.	WR	fixar	VERB	FIXAÇÃO	NOUN
8	L2-L2	unrel.	WR	verão	NOUN	PEQUENO	ADJ
8	L2-L2	unrel.	WR	difícil	ADJ	ESPERAR	VERB
8	L2-L2	unrel.	WR	capa	NOUN	SORVETE	NOUN
8	L2-L2	unrel.	WR	inverno	NOUN	FAMOSO	ADJ
8	L2-L2	unrel.	WR	passo	NOUN	DESENHO	VERB
8	L2-L2	unrel.	WR	roda	NOUN	COELHOS	NOUN
8	L2-L2	unrel.	WR	anual	ADJ	PREGUIÇA	NOUN
8	L2-L2	unrel.	WR	horrível	ADJ	TECIDO	NOUN
8	L2-Hung		WR	peçoal	ADJ	JOBBRA	
8	L2-Hung		WR	estudo	NOUN	KERESNI	
8	L2-Hung		WR	lixo	NOUN	KÉSZÜLT	
8	L2-Hung		WR	festa	NOUN	KIDOBVA	
8	L2-Hung		WR	missão	NOUN	KIFEJLÉS	
8	L2-Hung		WR	cópia	NOUN	KINYOMJA	
8	L2-Hung		WR	anel	NOUN	KIRÚGÁS	
8	L2-Hung		WR	resto	NOUN	KITÖMÖTT	
8	L2-Hung		WR	dança	NOUN	KORHOL	
8	L2-Hung		WR	ladão	NOUN	KÖZELI	
8	L2-Hung		WR	nota	NOUN	KURUZSLÓ	
8	L2-Hung		WR	tonto	ADJ	KUTATÓ	
8	L2-Hung		WR	selva	NOUN	KUTATUNK	
8	L2-Hung		WR	raso	ADJ	KUTYAGOL	
8	L2-Hung		WR	cela	NOUN	LÁBAZAT	
8	L2-Hung		WR	bagunça	NOUN	ZUHANÁS	

APPENDIX 2 – Suffix distribution

L1	suffix	n.		L2	suffix	n.
	IG	8			OR	2
	LOOS	3			IA	2
	LIJK	7			OSO	4
	ER	2			ADA	3
	ERIJ	2			EZ	2
	BAAR	3			ADO	3
	ERIG	3			EZA	2
	HEID	2			ANTE	2
	ZAAM	3			ENTO	1
Total	9 types				ÃO	2
					MENTE	2
					AL	2
					URA	3
					IDÃO	2
				total	14 types	

APPENDIX 3 – Autobiographical information about the participants

no. part.	age	AoA	total period	learning setting	context of use	stay in Brazil	education	other languages		
1	39	26	13	formal(1/2yr)	home/fam/social/work	9	superior	English	Spanish (interm)	
2	45	32	13	Contact	home/fam/social/work	9	superior	English	French (intermed)	
								Alemão (interm),		Spanish
3	28	23	5	formal (1 yr)	work/social/home(4 yrs)	4	superior	French,	Spanish,	English
4	39	21	18	formal (1/2yrs)	work/social	8	superior	English,	Spanish (interm)	
5	36	25	11	contact	work(50%)/social(80%)	9	superior	English,	French,	Spanish
								German(intermed),		
6	33	24	9	contact	home/fam/social/work	6	superior	Spanish,	English,	French
7	41	29	12	contact	home/fam/social/work	5	superior	English,	Spanish,	alemao
8	25	21	4	contact	fam/work(50%)/social	4	superior	French,	German,	English
9	32	25	7	formal(1/2yrs)	home/fam/social/work	7	superior	English		
10	34	23	11	formal(2yr)	home/fam/social/work	4	superior	English		
11	48	28	20	contact	work/social(50%)/co	9	cient.prep.	German,	French,	English
12	33	23	10	contact	home/work/social	4,5	superior	English,	German,	
13	26	21	5	contact/self study	home/work/social	4	superior	English	French(intermed)	
14	35	21	14	contato		7	superior	Spanish,	English,	
15	42	31	11	formal(2yrs)	home/work/fam/social	4	superior	German,	French,	English
16	35	22	13	formal(1/2yr)	home/fam/social/work	9	superior	English,	Spanish	
17	29	24	5	formal	home/fam/social/work	4	superior	English,	French,	alemao
18	41	26	15	contact	home/fam/social/work	15	cient.prep.	English,	Spanish	(intermed)

APPENDIX 4 –Proficiency judgment

4.1 Questionnaire

Pronunciation:

A. On a scale from 1 to 5, how would you rate the participants' speech:

1. you almost do not notice he/she is a foreigner, only hearing after a few sentence, paying a lot of attention;
2. you notice he or she is a foreigner, but she/he sounds like us;
3. Portuguese can be understood well, but the foreign accent is very clear;
4. The accent is so strong that it sometimes hinders comprehension of a few words;
5. The accent stops me from understanding.

B. The accent:

1. Marks only a few specific;
2. Marks the whole speech;
3. Gets in the way of communication.

Vocabulary:

C. The participant used:

1. a wide variety of words;
2. simple but adequate words;
3. words that were not adequate for the context.

D. The participant

1. has vocabulary to name specific sites and objects in a room;
2. has vocabulary to name objects and action, and when necessary, paraphrases;
3. has a limited vocabulary, repeating the same words, and paraphrasing, occasionally using Dutch.

Fluency:

E. The participant's speech:

1. flows freely, resulting in clear comprehension;
2. generally flows, with a few pauses during which the participants tries to structure the message;
3. does not flow, many pauses and hesitation to structure the message, resulting in difficult comprehension for the listener.

Structure:

F. Sentence structure:

1. is complete;
2. is usually complete, sometimes a preposition or some connection is missing;
3. Is broken, sometimes verb subject or complement, is absent from the language.

G. The participant used correct verbal inflection:

1. varies;
2. simple, with a little variation;
3. simple, with a little variation, with some errors;
4. simple, with many errors;
5. practically did not carry out verbal inflection.

H. The participant carries out agreement:

1. without any mistakes;
2. with a few mistakes;
3. made same constant errors, but he/she corrected himself;
4. got more wrong than right;
5. did not carry out any agreement.

4.2 Assessment

participant	pronunciation	vocabulary	fluency	structure
	<i>points/total</i>			
1	2,7/3	2,8/3	2,7/3	4,6/5
2	2,8/3	2,2/3	2,3/3	3,7/5
3	2,8/3	2,3/3	2,7/3	4,9/5
4	2,9/3	2,5/3	2,3/3	4,6/5
5	2/3	1,7/3	2/3	2,2/5
6	2,8/3	2,8/3	2,7/3	4,9/5
7	1,8/3	1,5/3	1,5/3	2,9/5
8	1,9/3	2,4/3	2/3	3,3/5
9	1,9/3	2,3/3	2,3/3	4,4/5
10	2,5/3	2,5/3	2,3/3	4,2/5
11	2,1/3	2,7/3	2/3	4,2/5
12	1,8/3	2,3/3	2/3	3,7/5
13	2/3	2/3	1,7/3	3,3/5
14	2/3	2/3	2/3	3,3/5
15	2/3	1,8/3	1,7	3/5
16	2,6/3	2,7/3	2,7/3	4,2/5
17	3/3	2,8/3	2,3/3	4,6/5

O teste de Cloze como instrumento de medida da proficiência em leitura: fatores linguísticos e não linguísticos

Cloze test as an instrument to measure reading proficiency: linguistic and non-linguistic factors

Kátia Nazareth Moura de Abreu

Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
kabreu00@gmail.com

Daniela Cid de Garcia

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
cid.daniela@gmail.com

Katharine de Freitas P. N. A. da Hora

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
freitas.katharine@gmail.com

Cristiane Ramos de Souza

Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro / Brasil
tianeramoss@gmail.com

Resumo: Este artigo tem como objetivos apresentar uma revisão da literatura do uso do teste de Cloze como ferramenta para medir o desempenho em leitura, discutindo seus limites e possibilidades, verificar a sensibilidade do teste para variáveis de natureza linguística e não linguística, e apresentar uma agenda de pesquisa que considere o uso de desenhos qualitativos originais para acessar diferentes aspectos da leitura. Para verificar a sensibilidade do teste, manipularam-se as variáveis tipo de palavra (lexical e funcional), tamanho de texto (curto e longo) e experiência leitora (medida em termos de nível de escolaridade). Observou-se o desempenho de alunos do Ensino Fundamental e do Ensino Médio. Os resultados demonstraram que os alunos do Ensino

Fundamental apresentaram mais dificuldade com palavras funcionais e com os textos longos, indicando que o teste é sensível a fatores linguísticos e não linguísticos. A partir dos resultados e das análises, discute-se como esse teste pode ser útil para avaliar diferentes processos cognitivos envolvidos na leitura.

Palavras-chave: compreensão leitora; teste de Cloze; ensino; palavras lexicais; palavras funcionais.

Abstract: The aims of this article are to present a review of the literature concerning the use of the Cloze test as a tool to measure reading performance, discussing its limits and possibilities, access the sensibility of the test for linguistic and non linguistic variables, and present a research agenda that considers the use of original qualitative designs to access different aspects of reading. To verify the sensibility of the test, we manipulated the variables word type (function words and content words), text length (short and long) and reading experience (measured in terms of level of schooling). We observed the performance of students from Elementary level and from High School. The results show that Elementary-level students had more difficulty with function words and with long texts, which indicates that the test is sensible to linguistic and non-linguistic factors. From the results and analyses, we draw a discussion of how the cloze test can be useful to evaluate different cognitive processes involved in reading.

Keywords: reading comprehension; Cloze test; education; content words; function words.

Recebido em: 5 de janeiro de 2017

Aprovado em: 10 de abril de 2017

1 Introdução

A compreensão de um texto pode variar de acordo com as circunstâncias de leitura e depender de fatores complexos presentes não só na superfície do texto como também na ativação, por parte do leitor, de um conjunto de saberes implícitos nesse ato de ler. A leitura exige processos de percepção, de cognição, assim como exige conhecimento da língua, da gramática, do assunto e conhecimento sobre como ler. Além

disso, envolve uma combinação de informação visual e não visual. Ela é uma interação entre o leitor e o texto (SMITH, 2003).

O objetivo deste artigo é apresentar uma revisão da literatura do uso do teste de Cloze, discutir os alcances do teste como medida da compreensão leitora, verificar como ele pode ser sensível a fatores linguísticos e não linguísticos – uma vez que a leitura é um amálgama de fatores de diferentes naturezas –, e apresentar uma agenda de como esses testes podem ser qualitativamente desenhados de forma a acessar faculdades distintas envolvidas no ato de ler.

O teste de Cloze é uma técnica que se compõe por um texto em que se omitem todas as quintas palavras que são substituídas por um traço de tamanho similar ao da palavra omitida. Essa técnica vem sendo utilizada ao longo dos últimos quarenta anos no Brasil como um instrumento de diagnóstico da compreensão em leitura – em que avalia a compreensão em leitura, interpreta os resultados e classifica os participantes em níveis de desempenho – e como um instrumento de intervenção – em que desenvolve a habilidade da compreensão em leitura ao conceder ao participante a possibilidade de prever as palavras omitidas pelo uso de seu conhecimento prévio e pelo domínio das estruturas linguísticas.

De modo a avaliar o impacto dos fatores tipo de palavra, extensão textual e escolaridade no desempenho de leitores, realizamos um teste de Cloze com estudantes do 6º ano do Ensino Fundamental e do 1º ano do Ensino Médio. As variáveis foram escolhidas com a intenção de incluir elementos do conhecimento linguístico (diferença entre palavras funcionais e lexicais) e elementos relacionados a processamentos cognitivos mais gerais, como atenção, memória e motivação. Além disso, incluímos grupos de diferentes escolaridades para verificar se a experiência leitora teria impacto no desempenho.

Com relação ao fator linguístico tipo de palavra, é sabido que, do ponto de vista do processamento da informação, há uma diferenciação de valor entre palavras lexicais e palavras funcionais no que tange à contribuição para recuperar o sentido do *input* (VAN PATTEN, 1996). Palavras lexicais apresentam maior conteúdo semântico, ao passo que palavras funcionais carregam sobretudo informação gramatical, de modo que estudos em psicolinguística tendem a estabelecer uma diferença de representação e processamento entre essas duas categorias. Palavras lexicais tendem ainda a ser maiores em tamanho e representam

um inventário aberto, constituindo um conjunto maior do que o das palavras funcionais. Sendo assim, a competição entre palavras lexicais possíveis em uma posição na sentença tende a ser, na maioria dos casos, consideravelmente maior. Do ponto de vista do acesso lexical, estudos mostram que existe maior facilidade de processamento para palavras funcionais. Essa diferença é atribuída a fatores como previsibilidade e familiaridade da palavra, podendo ser afetada pela frequência (SEGALOWITZ; LANE, 2000) e pelo fato de a palavra ser apresentada isoladamente ou em um contexto significativo (TAFT, 1990). Estudos de aquisição revelam que as crianças começam o processo de produção com palavras lexicais, sendo esse tipo de palavra mais saliente. Por outro lado, quando se trata da percepção, estudos revelam que crianças são sensíveis à presença de palavras funcionais e apresentam expectativas quanto à distribuição dessas palavras na sentença (VAN HEUGTEN; CHRISTOPHE, 2015). À medida que crianças vão adquirindo palavras funcionais, elas vão estabelecendo padrões distribucionais de coocorrência com a categoria de palavras lexicais. A aquisição de palavras funcionais tem, portanto, papel relevante na aquisição da linguagem, contribuindo no processo de reconhecimento das categorias gramaticais da língua (BRAINE, 1987; BRUSINI *et al.*, 2017, 2016). Há discussão na literatura quanto a se essa contribuição viria devido a uma análise distributiva linear, ou a uma relação estrutural hierárquica entre os elementos das frases. Esses estudos em conjunto atestam a natureza distinta entre palavras funcionais e palavras lexicais, com impacto na aquisição da linguagem e no processamento adulto.

De acordo com estudos sociolinguísticos, as variáveis linguísticas e não linguísticas agem de maneira integrada. Tais estudos apontam que as variáveis não linguísticas podem favorecer ou inibir o emprego de determinada forma linguística. A variável nível de escolaridade, uma das variáveis não linguísticas de que trata este estudo, vem sendo amplamente testada e estudada a fim de observar se o seu grau gera alguma influência sobre os falantes quanto à utilização da norma culta padrão do português brasileiro. Gonzales *et al.* (2007) afirmam que a variável escolaridade tem influência sobre o uso de diferentes fenômenos linguísticos e que tal influência pode estar correlacionada com a função social que a escola exerce.

É sabido que

a escola gera mudanças na fala e na escrita das pessoas que as frequentam e das comunidades discursivas. Constata-se, por outro lado, que ela atua como preservadora de formas de prestígio, face a tendências de mudança em curso nessas comunidades (VOTRE, 2012, p. 51).

Pereira e Araújo (2016) fizeram um levantamento acerca da influência da variável escolaridade sobre a variação da concordância verbal com sujeito na 3ª pessoa do plural no português brasileiro. As autoras perceberam que a maior parte dos estudos mostra que o uso da variante de prestígio aparece com maior frequência em registros de indivíduos com grau de escolaridade maior do que daqueles que possuem pouca escolaridade. As autoras afirmam, em conformidade com Cyranka e Pernambuco (2008), que a privação de uma educação formal impossibilita o indivíduo de usar certas formas linguísticas de prestígio e, conseqüentemente, de transitar em diversas situações comunicativas em que essas formas são empregadas.

Parente *et al.* (2009), em um estudo sobre as evidências do papel da escolaridade na organização cerebral, observaram que há mudanças comportamentais, de desempenho em avaliações neuropsicológicas, e estrutural-funcionais, verificadas em exames de neuroimagem da estrutura e da funcionalidade do cérebro humano de acordo com os níveis de escolaridade dos indivíduos testados. Para as autoras, a variável escolaridade mostra-se como uma variável multidimensional, pois a ela são somados fatores como hábitos de leitura e escrita cultivados pelos indivíduos por interesses pessoais, estilo de vida ou por necessidades profissionais.

Outro fator não linguístico que manipulamos neste estudo foi o tamanho do texto. De acordo com Kato ([1986], 2001), extrair o significado de um enunciado depende dos elementos presentes na memória de um indivíduo e no modo como a memória funciona. Ao ler um texto, um indivíduo ativa diferentes tipos de memória – de curto, médio e longo prazo. A relação entre o tipo de memória ativado e o conteúdo do texto determina se a estratégia de leitura será predominantemente descendente (*top-down*) – quando o leitor parte do seu conhecimento prévio para construir o sentido do texto – ou ascendente (*bottom-up*) – em que o leitor reconhece palavra por palavra, juntando-as para construir unidades maiores e gerando o sentido a partir da combinação dessas unidades.

Há pesquisas que afirmam que textos curtos tendem a favorecer uma leitura do tipo ascendente, uma vez que a pouca demanda de memória permitiria maior atenção a detalhes e encorajaria estratégias de processamento palavra por palavra (KINTSCH; VAN DIJK, 1978; SWAFFAR; ARENS; BYRNES, 1991). Além disso, textos curtos também são vistos como potencialmente mais difíceis pela quantidade reduzida de informações, ao passo que textos longos seriam mais facilmente compreendidos, por fornecerem mais conteúdo por meio do qual se pode apreender sentido. Por outro lado, há de se considerar o fator atenção quando se examina o papel da extensão textual na compreensão leitora e no desempenho de estudantes em testes de leitura. A experiência de professores em sala de aula aponta para maior dificuldade de concentração e para um desempenho comprometido na leitura de textos longos em comparação com textos curtos. Essa constatação é legitimada pela preponderância de textos relativamente curtos em livros didáticos.

Na próxima seção, apresentamos uma revisão da literatura do teste de Cloze como ferramenta de medida da compreensão leitora. Em seguida, apresentamos um experimento que manipula quantitativamente as variáveis de ordem linguística e não linguística elencadas acima, verificando a sensibilidade do teste a essas variáveis. Por fim, conduzimos uma discussão sobre como essas variáveis linguísticas e não linguísticas podem interagir com a proficiência em leitura, propondo uma agenda de pesquisa para futuros trabalhos, que pode se beneficiar de uma abordagem qualitativa na extração dos itens nos testes de Cloze.

2 O teste de Cloze

O termo *Cloze* é derivado da palavra inglesa *closure* que significa *fechamento* (COLLINS, 2009) e que, por sua vez, dá nome a um princípio da Psicologia da Gestalt. Esse termo foi utilizado por Wilson L. Taylor em 1953 para nomear um procedimento que visava à mensuração da leituraabilidade do texto (*readability*), cuja base está fundamentada na Teoria da Informação, na Psicologia da Gestalt e na noção estatística de amostra aleatória. Em artigo seminal, Taylor (1953, p. 416 *apud* ADELBERG; RAZEK, 1984, p. 111) definiu o *Cloze* como

um método de interceptar uma mensagem de um emissor (escritor ou falante), suprimindo, pelo apagamento de partes, seu padrão de língua, e então apresentando essa

mensagem aos receptores (leitores ou ouvintes) para que as suas tentativas de reconstrução do padrão potencialmente rendam um número considerável de unidades de Cloze (tradução nossa).

Taylor (1953, p. 416 *apud* ADELBERG; RAZEK, 1984, p. 111) segue explicando que uma unidade de Cloze é

qualquer ocorrência de uma tentativa bem sucedida para reproduzir acertadamente uma parte suprimida de uma 'mensagem'... pela decisão, a partir do contexto que permanece, de qual deveria ser a parte apagada (tradução nossa).

Assim, em um experimento com a técnica de Cloze é possível visualizar a Teoria da Informação na ideia de que o emissor e o receptor devem compartilhar o mesmo conjunto de mensagens possíveis, a Psicologia da Gestalt na ideia de se ter o leitor reconstruindo uma mensagem e alcançando a sua compreensão, pelo preenchimento de palavras anteriormente suprimidas, e a noção estatística de amostra aleatória na ideia de que diferentes classes de palavras ganham representatividade proporcional nas lacunas que são abertas no texto. Ou seja, a probabilidade de escolha de palavras lexicais, por exemplo, seria proporcional à escolha de palavras funcionais. Sobre esse aspecto, Adelberg e Razek (1984, p. 112) comentam que o teste de Cloze parece permitir a medida tanto de um componente lexical quanto de um componente funcional, dependendo do tipo de palavra selecionada para ser suprimida. Já Condemarin e Milicic (1988) ponderam que apesar de Taylor ter lançado mão da Psicologia da Gestalt para fundamentar a técnica de Cloze, ela também é considerada um instrumento cognitivo pelos teóricos cognitivistas, pois esses assumem que tal técnica lida com processos cognitivos como percepção e memória e veem no *Cloze* um processo interativo entre o leitor e o texto.

O teste de Cloze, em seu formato original, prevê a seleção de um texto com cerca de 250 palavras. A cada cinco palavras suprime-se a quinta sistematicamente. No lugar dessas palavras são colocados espaços em branco proporcionais aos tamanhos das palavras. Um grande número de estudos, porém, tem utilizado formas metodológicas variadas de construção do teste de Cloze. Uma das propostas de modificação da versão original é que já conseguiu formar um grupo significativo de

pesquisadores que lhe são favoráveis está baseada no pressuposto de que, para se testar aspectos específicos da compreensão textual, elementos precisos devem ser suprimidos. Essa proposta, segundo Sadeghi (2014, p. 1) já promoveu a variedade de *Cloze* chamada de “*Cloze* racional” que se contrapõe a “*Cloze* randômico” (tradução nossa). No teste de *Cloze* racional, as omissões são feitas com base em critérios determinados (não mecânicos) enquanto no de *Cloze* randômico essas são feitas aleatoriamente (mecânicos).

Os adeptos do *Cloze* racional (ABRAHAM; CHAPELLE, 1992; SADEGHI, 2014) argumentam que se o objetivo do teste é medir a compreensão do texto, a ideia da supressão randomizada precisa ser descartada em favor da supressão racional. É possível perceber, assim, que, dependendo do propósito da aplicação do teste de *Cloze*, tipos diferentes de *Cloze* racionais podem ser construídos, como, por exemplo, o *Cloze* lexical – em que se omitem os itens lexicais como substantivos, adjetivos, verbos e advérbios – e o *Cloze* gramatical – em que se omitem os itens funcionais como preposições, artigos, conjunções (Santos *et al.*, 2009).

Outra proposta de modificação do *Cloze* está relacionada ao tipo de resposta. Na versão clássica, as respostas são abertas, ou seja, o participante pode escolher qual a palavra que preenche a lacuna corretamente. Em versões modificadas, as respostas são fechadas, ou seja, o participante terá de escolher, entre as alternativas oferecidas, a que preenche a lacuna corretamente. Nesse caso, temos o *Cloze* múltipla escolha e suas variações como o *Cloze* labirinto, o *Cloze* pareado e o *Cloze* restringido (OLIVEIRA *et al.*, 2009). Os defensores dessa proposta argumentam que o uso de respostas fechadas torna o teste mais prático em termos de pontuação.

Por fim, observamos a variação existente no modelo de aplicação que pode ser uma aplicação coletiva – como quando o teste de *Cloze* é utilizado como ferramenta pedagógica por professores em uma turma ou mais turmas – ou uma aplicação individual – como quando o teste de *Cloze* é utilizado como instrumento de diagnóstico por psicólogos na avaliação de um paciente. Oliveira *et al.* (2009) nos mostram que o *Cloze* pós-leitura oral e o *Cloze* interativo são exemplos de modelo de aplicação individual. A escolha de um modelo de aplicação vai sempre depender da gênese do estudo e do profissional que vai lançar mão dele.

Segundo Santos *et al.* (2007), o procedimento de *Cloze* possui duas formas de correção: a correção literal e a correção sinônima.

A correção literal considera como acerto o preenchimento correto da palavra exata que foi omitida, respeitando inclusive grafia e acentuação gráfica. No caso da correção sinônima, considera-se como correto o preenchimento da lacuna, não somente com a palavra que foi excluída, mas também com um sinônimo da palavra omitida. (SANTOS *et al.*, 2007, p. 42)

As mesmas autoras, porém, em artigo de 2009, apresentam um terceiro tipo de correção, a correção ponderada. Nessa forma, o avaliador faz uma ponderação sobre a palavra que foi escrita pelo participante, verificando se a palavra utilizada se aproxima ou não do esperado e não se preocupa com correção ortográfica (OLIVEIRA *et al.*, 2009).

Independentemente do tipo de correção adotada, no que diz respeito à pontuação, em geral, é atribuído um ponto a cada acerto. A pontuação final do participante é obtida pela soma dos pontos, o que pode variar de 0 a 25 pontos ou de 0 a 50 pontos, ou ainda de 0 a “n” pontos, dependendo da quantidade de lacunas existentes no texto. Entretanto, Santos *et al.* (2002, p.550) nos alertam que

[e]sta interdependência entre os escores e a dificuldade dos testes é uma das limitações dos modelos da psicometria clássica, na qual os escores são calculados atribuindo o mesmo peso para itens heterogêneos quanto à complexidade.

Por esse motivo, outra forma de pontuação mensura o nível de dificuldade do item. Nesse caso, são atribuídos valores diferentes para cada item de acordo com o seu grau de complexidade. Esses valores são estabelecidos por especialistas em língua portuguesa chamados de ‘juízes’, os quais, após exame dos itens suprimidos, indicam os níveis de dificuldade, a saber: fácil, médio, moderado e difícil. Assim, presencia-se uma variabilidade na soma dos pontos, pois os itens classificados como fáceis recebem menos pontos ao contrário dos itens classificados como difíceis que recebem mais pontos.

Quanto à interpretação da pontuação alcançada pelo participante, podem ser adotadas análises estatísticas a fim de se verificar a influência de variáveis e a significância ou não dos resultados estatísticos. Outra opção é adotar o parâmetro de classificação proposto por Bormuth (1968) que prevê dois níveis de leitura: o nível instrucional e o nível

independente. O nível instrucional correspondente ao percentual de acerto na faixa de 45% a 57% do texto indica que o leitor alcançou compreensão suficiente, mas que ainda há necessidade de auxílio adicional externo (do professor, do livro, por exemplo). O nível independente corresponde ao percentual de acerto superior a 57% do texto e indica que o leitor tem autonomia de compreensão em leitura. Santos *et al* (2002) acrescentam, ainda, um nível que denominam de frustração, que abrange o percentual de acerto de até 44% do total do texto e indica que o leitor extraiu poucas informações da leitura, ou seja, alcançou pouco êxito na compreensão. Para a classificação proposta por Bormuth (1968), encontramos também uma observação em Santos *et al.* (2002, p. 550) afirmando que essa “aplicação para novos textos só pode ser feita se esses textos tiverem um nível de dificuldade semelhante ao dos textos que Bormuth utilizou em seu estudo”.

Então, pela descrição apresentada aqui sobre o *Cloze*, podemos concluir que esse é um tipo de experimento econômico na preparação, de simples aplicação (lápiz e papel), porém é um experimento que exige cuidados na seleção dos textos, na abertura das lacunas e na correção/pontuação.

Quanto a isso, Suehiro (2013, p. 224) alerta que

No que se refere à estruturação do texto, além dos parâmetros de omissão das palavras, é preciso considerar também a questão de sua dificuldade e dos itens ou palavras omitidas. Quanto à dificuldade, verifica-se que o assunto abordado no texto também é um aspecto que deve ser considerado, tendo em vista que foram encontradas diferenças de desempenho com relação a essa variável; logo, sob essa perspectiva, a familiaridade com o assunto aumenta a probabilidade de acerto das respostas dadas, o que introduz uma variável interveniente na comparação dos desempenhos obtidos pelos indivíduos durante a realização da técnica (COHEN, 1975; OLIVEIRA; BORUCHOVITCH; SANTOS, 2009; PAGE, 1975).

Neste estudo, realizamos um teste de *Cloze* cujo objetivo foi o de investigar se o processamento de um texto durante a leitura pode ser modulado por um fator de ordem linguística, como a diferença entre palavra lexical e palavra funcional, e por um fator de ordem não

linguística, como o tamanho do texto. O tamanho do texto pode impactar fatores de processamento como atenção e memória de trabalho. O tipo de palavra tem por fim constatar o papel de fatores linguísticos que podem ser intrínsecos ao item ou relacionados a sua distribuição em relação a outros itens na construção do significado de uma proposição. Além disso, verificamos como esses fatores podem ser impactados pelo nível de ensino. Para tal fim, realizamos o teste com dois grupos distintos: alunos do 6º ano do Ensino Fundamental e alunos do 1º ano do Ensino Médio, já que estudos apontam que o nível de escolaridade/idade é um fator a ser considerado em relação ao saber linguístico.

3 O uso do Teste de Cloze

Com amplo uso na esfera educacional, o teste de *Cloze* é uma ferramenta de diagnóstico da compreensão de leitura. Essa ferramenta já foi utilizada por diferentes pesquisadores, tanto em abordagens básicas sobre o processamento cognitivo de sentenças, quanto em pesquisas de dimensão mais aplicada, com implicações para práticas de ensino de língua (FEDERMEIER *et al.*, 2007; SANTOS, 2004; BALIZA; SILVA, 2015). A princípio, pode-se pensar que essa ferramenta esteja limitada à verificação do vocabulário, mas o teste pode ser também utilizado em frases soltas, em que a verificação é mais local, ou em textos, em que a verificação é macroestrutural.

Desde que foi criado, o *Cloze* tem sido considerado um recurso confiável para avaliar a compreensão de textos escritos e tem sido utilizado em uma grande variedade de pesquisas sobre o estudo da leiturabilidade, bem como em estudos que buscam avaliar a influência da posição sintática das palavras na sentença. Além dessas aplicações, foi também utilizado em exames de proficiência em língua estrangeira e em exames de seleção de diversas universidades (CUNHA, 2009). Atualmente, esse teste vem sendo utilizado por pesquisadores de leitura e de linguagem, em uma perspectiva multidisciplinar, que contempla áreas como Psicologia, Comunicação, Fonoaudiologia, Linguística e Educação e apresenta duas vertentes principais: a que focaliza a avaliação de mecanismos psicolinguísticos que atuam na compreensão da língua escrita e a que focaliza o estabelecimento de um diagnóstico das dificuldades associadas à compreensão da língua escrita (SANTOS *et al.*, 2009).

Em um estudo feito por Baliza e Silva (2015) utilizando a técnica de *Cloze*, foi observada, entre outros fatores, a influência do nível de escolaridade na compreensão da leitura. O teste foi aplicado em alunos de 7ª e 8ª séries. Os autores observaram que os alunos com mais escolaridade compreendem melhor o que leem em relação aos demais alunos.

Há ainda estudos que combinam a técnica de *Cloze* com a percepção de desempenho escolar. Oliveira, Boruchovitch e Santos (2009) investigam a relação entre compreensão em leitura, desempenho escolar real e desempenho escolar autopercebido de alunos da 5ª série do ensino fundamental e verificam que a compreensão em leitura encontra-se em nível básico. Demonstram que, apesar de os alunos conseguirem, com esse nível, a habilidade mínima para a aprendizagem, eles não alcançam o nível de proficiência em leitura. Quanto à relação entre desempenho escolar em Língua Portuguesa e compreensão em leitura, as autoras constataam que os alunos com melhor aproveitamento na disciplina também exibem maior compreensão do texto. Revelam, ainda, que os alunos com autopercepção negativa do desempenho obtiveram menor pontuação no teste de *Cloze* e que o mesmo fato ocorrendo inversamente, pode sugerir que os alunos estão cientes de seu próprio desempenho escolar. Por fim, apontam a necessidade de que propostas de intervenção sejam adotadas para a elaboração do diagnóstico e da remediação de dificuldades específicas da compreensão em leitura.

Igualmente, tendo como foco de investigação a compreensão leitora no ensino fundamental, Joly e Piovezan (2012) empreenderam um estudo com o objetivo de avaliar o Programa Informatizado de Leitura Estratégica (PILE) desenvolvido por Joly (2008) que é um programa destinado a promover a compreensão em leitura de alunos do 5º ao 9º ano do ensino fundamental. O conteúdo do programa é formado por uma série de trechos de histórias em que foram aplicados o Sistema Orientado de *Cloze* (JOLY, 2006, 2009), em que as dificuldades relacionadas ao número de vocábulos que aparecem entre as lacunas e as pistas adicionais como tamanho do espaço e banco de palavras surgem em sequência gradativa, tendo como critério de organização dos textos, por nível de dificuldade, o Sistema Orientado de *Cloze* (SOC). Os resultados obtidos não só comprovam a eficácia do Programa como também apontam para a importância da proposta interventiva com a finalidade de incrementar o desempenho em compreensão leitora. Além disso, corroboram a eficiência do Sistema Orientado de *Cloze* em programas de intervenção

(JOLY, 2007, 2009) e o procedimento de Cloze como técnica de diagnóstico e de remediação, conforme já demonstrado por Santos *et al.* (2002), Santos (2004) e Cunha (2009).

Joly *et al.* (2014), dando continuidade a pesquisas sobre a compreensão em leitura, realizaram um estudo de avaliação da compreensão leitora em alunos do ensino fundamental do 6º ao 9º ano com a utilização do Sistema Orientado de Cloze (SOC). Nesse estudo, foram consideradas variáveis como idade, gênero e escolaridade a fim de verificar se havia diferenças na compreensão em leitura. A análise dos resultados indicou uma diferença significativa entre os gêneros – em que as meninas apresentaram melhor desempenho – e, entre o ano de escolaridade – em que mais uma vez os alunos de séries mais adiantadas apresentaram melhor desempenho. Quanto à idade, porém, os resultados não se mostraram significativos.

Suehiro e Boruchovitch (2016) também realizaram um estudo direcionado a alunos do ensino fundamental, especificamente com alunos do 6º ao 9º ano, e consideraram como variáveis o gênero, a escolaridade e a idade. Após a análise, os resultados revelaram que os participantes se classificariam em um nível básico de compreensão em leitura que não era compatível com o seu grau de escolarização. Assim como os dados de Joly *et al.* (2014) houve diferença significativa entre os gêneros com um melhor desempenho das meninas e, do mesmo modo, com relação ao nível de escolaridade em que os alunos de anos mais avançados apresentaram melhor desempenho. A discrepância entre os estudos se deu por conta da variável idade, pois em Joly *et al.* (2014) esse aspecto não se configurou como significativo ao contrário da descrição de Suehiro e Boruchovitch (2016) em que os mais velhos demonstraram melhor desempenho.

Em referência ao ensino médio, o estudo de Oliveira, Cantalice e Freitas (2009), apresentou objetivos semelhantes aos estudos anteriormente apresentados, porém com um objetivo específico de explorar os acertos relativos aos itens. A questão dos itens é particularmente importante no teste de Cloze porque a acurácia no preenchimento do item omitido possivelmente não depende somente da habilidade de leitura dos participantes como pode também depender da dificuldade do item a ser preenchido.

Assim, nesse estudo, os itens eram compostos por oito classes de palavras (pronomes, adjetivos, advérbios, artigos, conjunções, preposições, substantivos, verbos) e por uma subclasse (locução adjetiva) e foram

classificados como fáceis, de dificuldade média e difíceis. Os resultados indicaram que os alunos possivelmente apresentam dificuldades para ler e compreender o que leem e que essa compreensão em leitura encontra-se abaixo do nível básico, fato que não era esperado nesse nível de escolaridade. As autoras ressaltaram que esses resultados são exploratórios e apontaram que com relação às dificuldades dos itens, os resultados mostraram que as categorias com maior índice de acertos foram as preposições e os substantivos e a com menor índice, a locução adjetiva. As autoras ressaltam que esses dados são exploratórios e apontam para a conveniência de se realizar uma análise mais precisa.

Em consonância com o que se pretende apresentar aqui, os estudos referenciados acima consideraram variáveis linguísticas e variáveis não linguísticas. Cabe ressaltar, no entanto, que, se por um lado, o conjunto de trabalhos reportados está longe de compor uma lista exaustiva de estudos com Cloze nos ensinos fundamental e médio, por outro lado, ilustram a aceitabilidade e a credibilidade alcançada por esse procedimento psicolinguístico no meio acadêmico.

A relevância da técnica de Cloze reside na contribuição para a investigação do problema da compreensão e do desempenho em leitura em todos os níveis de ensino. Sua utilidade pode ser observada pelo número de publicações que tratam de seu emprego seja para traçar diagnóstico de compreensão em leitura seja para a elaboração de programas de intervenção e remediação, como atestam os dados de Suehiro (2013) ao descrever a produção científica sobre o teste de Cloze em periódicos de psicologia entre os anos de 2002 a 2012.

4 Acessando a sensibilidade do teste de Cloze a fatores linguísticos e não linguísticos

Este experimento teve como objetivo verificar a sensibilidade do teste de Cloze a fatores linguísticos e não linguísticos. Para tal fim, manipulamos as variáveis tipo de palavra (funcional e lexical), tamanho do texto e escolaridade dos participantes. O procedimento adotado na criação das lacunas foi o racional (ABRAHAM; CHAPELLE, 1992; SADEGHI, 2014) em que manipulamos o tipo de palavra retirada (funcional ou lexical). Assim, em um texto haverá maior quantidade de lacunas para o tipo de palavra lexical (foco) e menor quantidade de lacunas para o tipo de palavra funcional. Em outro texto, haverá o

contrário: maior quantidade de lacunas para as palavras funcionais (foco) e menor quantidade para as palavras lexicais.

Resumidamente, as variáveis independentes testadas foram (i) tipo de palavra, que se dividia entre palavras funcionais e palavras lexicais e (ii) tamanho do texto, que podia ser curto ou longo. Nesse teste tivemos apenas uma variável dependente: foram medidos os índices de erros no teste de Cloze. As variáveis independentes constituem um design 2x2 gerando as condições TLF – Texto Longo com palavra Funcional, TLL – Texto Longo com palavra Lexical, TCF – Texto Curto com palavra Funcional e TCL – Texto Curto com palavra Lexical (Tabela 1).

TABELA 1 – Design do experimento

6º ano Ensino Fundamental			1º ano Ensino Médio		
	Funcional	Lexical		Funcional	Lexical
Texto Longo	TLF	TLL	Texto Longo	TLF	TLL
Texto Curto	TCF	TCL	Texto Curto	TCF	TCL

A partir dos dados obtidos por este experimento, espera-se observar diferença significativa de desempenho entre os grupos, relacionada à extensão do texto e ao tipo de palavra. Com relação à extensão do texto, é esperado que leitores mais experientes sejam menos afetados pela dificuldade atencional gerada por textos longos. Com relação ao tipo de palavra, Cain e Nash (2011) apontam que a competência leitora de palavras funcionais é modulada pela idade dos leitores.

Considerando a variável tipo de palavra, isoladamente, esperamos encontrar dificuldade maior para o preenchimento de lacunas com palavras funcionais, já que, considerando fatores como distribucionalidade e coocorrência de palavras, deve ser mais custoso para o participante acessar uma palavra funcional a partir de uma palavra lexical. Isso se deve ao fato de as opções de palavras funcionais que podem ocorrer próximas a uma palavra lexical serem mais restritas que as opções de palavras lexicais próximas a uma palavra funcional. Sendo assim, ainda que a previsibilidade de uma palavra funcional seja maior em uma determinada lacuna em relação a uma palavra lexical, devido

ao repertório limitado do paradigma, é justamente esse número limitado que deve gerar a dificuldade. Como o protocolo utilizado no experimento aceita respostas sinônimas ou coerentes com o sentido do texto, é mais provável que o aluno conheça uma possibilidade dentro de um repertório maior do que dentro de um repertório menor.

Em relação à variável texto longo, isoladamente, a literatura sugere que o desempenho na compreensão leitora pode ser facilitado pela extensão do texto, uma vez que textos longos forneceriam mais informação e pistas para a interpretação. Por outro lado, acreditamos que o papel de fatores como atenção e motivação pode afetar significativamente o desempenho leitor, o que geraria a previsão de maior dificuldade na compreensão de textos mais longos, sobretudo no grupo de menor escolaridade.

Participantes

Participaram deste estudo 88 alunos de uma escola da rede Pública Estadual do Rio de Janeiro situada no município de São Gonçalo, com frequência regular, sendo 44 alunos do 6º ano do Ensino Fundamental e 44 alunos do 1º ano do Ensino Médio. A idade média dos participantes do 6º ano foi igual a 12 anos com desvio padrão de 1,38 e do ensino médio a média da idade foi igual a 17 anos com desvio padrão de 0,74.

Material

O material do teste foi composto por textos extraídos de livros didáticos, correspondentes às séries em que se encontram os participantes da pesquisa. Foram escolhidos textos de caráter ficcional, do gênero *crônica*, por ser um dos primeiros gêneros textuais com os quais os alunos entram em contato no Ensino Fundamental e por se considerar a indicação de Joly (2014, p. 239), quando afirma que

A técnica de Cloze com suas inúmeras possibilidades de variações favorece sua utilização com textos de diferentes conteúdos, sejam específicos e ligados a disciplinas escolares, sejam de caráter literário, como os utilizados no presente estudo.

Além disso, trata-se de um gênero mais cotidiano, confessional e que aborda assuntos corriqueiros. Com uma estrutura linguística próxima

da narração e com linguagem simples, esse gênero é tido como mais acessível não só a todo tipo de leitor, mas também, em geral, a todas as faixas etárias. Os textos selecionados para compor as versões do teste de *Cloze* foram três crônicas de Carlos Drummond de Andrade, intituladas “O assalto”, “No aeroporto” e “Recalcitrante”; e um relato de viagem, intitulado “O segredo de Mioko”, do autor Marco Clanflone. O critério utilizado na seleção desses textos também considerou o que está indicado por Coelho (2000) sobre a tipologia textual e sobre o gênero textual que estaria adequado aos alunos em conformidade com a faixa etária e com a série.

O texto “O assalto” com 218 palavras, apresentava 25 lacunas; o texto “No aeroporto” com 515 palavras apresentava 50 lacunas; o texto “Recalcitrante” com 629 palavras apresentava 50 lacunas e o texto “O segredo de Mioko”, com 394 palavras apresentava 30 lacunas. As lacunas foram intencionalmente abertas e o espaço demarcado era igual em todas as lacunas.

Procedimentos

Após contato inicial com a direção da escola em que houve a solicitação formal para a pesquisa, foi firmado o acordo de cooperação. A direção da escola disponibilizou os dias e os horários das turmas que seriam os mais adequados para a aplicação do experimento. Realizamos, então, a aplicação de forma coletiva, em sala de aula, conforme o agendamento previsto pela escola.

Antes de dar início ao teste, as pesquisadoras explicaram de forma clara todo o procedimento e tiraram as possíveis dúvidas sobre a tarefa a ser realizada. Os textos foram entregues aos alunos, em situação de sala de aula, pelas pesquisadoras. Os testes foram aplicados em diferentes momentos de um dia normal de atividades escolares, entre outras atividades previstas no planejamento de aula dos professores das diferentes turmas. Os participantes recebiam um texto contendo uma das quatro condições testadas (TLL, TLF, TCL, TCF), e tinham como tarefa preencher as lacunas com uma palavra que garantisse a coerência do texto. Além disso, os participantes responderam a questões de identificação, informando, na mesma folha do teste de *Cloze*, os seguintes itens: nome, idade e turma a qual pertenciam.

O teste foi realizado em ambiente tranquilo, bem arejado e iluminado. A realização da atividade teve duração média de 50 minutos por sala, não tendo sido estabelecido um limite de tempo para que os alunos concluíssem a tarefa.

5 Análise

Critérios de correção

No teste de *Cloze* podem ser adotados tradicionalmente dois critérios para a correção dos itens: o critério literal (ou objetivo) e o critério sinônimo (ou subjetivo). No primeiro, são considerados corretos os itens que correspondem às palavras usadas no texto original. Já no segundo critério, são considerados corretos os itens que garantem a coerência do conteúdo. Além desses, há o critério de ponderação (OLIVEIRA *et al.*, 2009) em que são considerados corretos os itens que se aproximam do ideal de preenchimento, não importando a correção ortográfica ou gramatical.

Neste trabalho, a avaliação do desempenho foi feita com base no critério ponderado e consistiu em atribuir um ponto para cada acerto, sendo consideradas corretas as palavras que completassem as frases com coerência. A fim de evitar um excesso de subjetividade na correção, as respostas foram apreciadas pelas pesquisadoras e houve a análise de juizes (FONSECA *et al.*, 2011) em que profissionais especializados em língua portuguesa examinaram e estabeleceram os níveis de dificuldades dos itens (OLIVEIRA *et al.*, 2009). O número de acertos máximo era de 50 (texto longo) ou 25/30 (texto curto), correspondentes ao número de omissões em cada versão do texto.

Análises dos dados

De início, foi feita a contabilização de erros e acertos de cada um dos indivíduos testados. Após a verificação dos dados de cada participante, foi necessário contrabalancear os dados, de modo que todas as condições fossem quantitativamente equivalentes para a análise estatística. Igualamos todos os dados a 50, conforme os dados do texto longo. Sendo assim, para um total de 50 lacunas, o número de erros do aluno 1, por exemplo, na condição TCF é 24, o que corresponde aos 12

erros de um total de 25. Tal procedimento foi feito com os dados de todos os alunos, tanto do Ensino Médio quanto do Ensino Fundamental, a fim de termos uma análise mais homogênea e mais confiável.

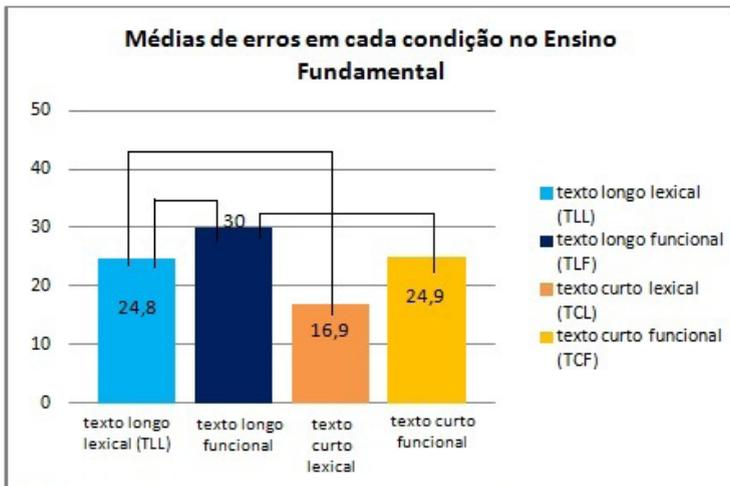
Após esse passo, fizemos as análises estatísticas usando o programa computacional R. Foram realizados dois testes: o teste T e a ANOVA. Fizemos dois tipos de análise: a análise *Within Subject*, em que comparamos os dados por grupo e a análise *Between Subject*, em que comparamos os dados entre os grupos testados.

6 Resultados

Como mencionado, o objetivo do trabalho foi verificar a sensibilidade do teste de Cloze aos fatores linguísticos tipo de texto (longo ou curto) e tipo de palavra (lexical ou funcional) e a fatores não linguísticos (escolaridade). O teste foi feito com alunos do sexto ano do Ensino Fundamental e do primeiro ano do Ensino Médio.

No Gráfico 1, em que são apresentadas as médias de erros no preenchimento da lacuna dos alunos do ensino fundamental, podemos observar que os alunos tiveram médias de erros bastante elevadas em todas as condições. Em uma comparação entre as condições testadas, é possível observar, visualmente, que os alunos apresentaram maior dificuldade ao preencher as lacunas da maneira esperada quando elas pediam uma palavra funcional e quando os textos eram longos. Em relação ao tipo de palavra (lexical ou funcional) os alunos apresentaram maior número de erros quando se tratava de palavras funcionais independentemente do tamanho do texto.

GRÁFICO 1 – Médias de erros no preenchimento da lacuna



Em uma análise estatística (T-teste) utilizando as médias de erros em cada condição observamos que a comparação entre as condições TLL e TCL é altamente significativa tendo $t(10)=9,58$ $p<0,0001^{***}$. Observamos também que nas condições TCL e TCF há significância sendo $t(10)=4,81$ $p<0,0007$. Na comparação entre as condições TLL e TLF também há significância sendo $t(10)=2,65$ $p<0,02$. Entre as condições TLF e TCF a significância mostrou-se mínima sendo $t(10)=3,03$ $p<0,01$.

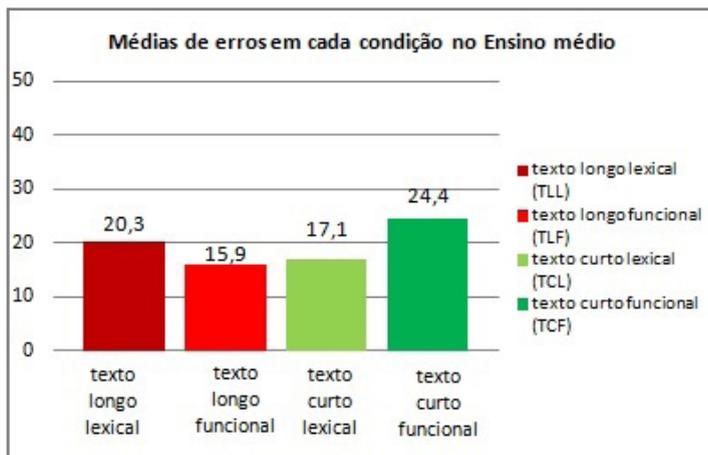
O teste de significância ANOVA mostrou-se altamente significativo para a variável Tipo de Texto $F(1,10) = 73,2$ $p<0,000007$ e também para a variável Tipo de Palavra $F(1,10)=20,5$ $p<0,001$. Portanto, quanto à análise *Whitin subjects* feita com os alunos do ensino fundamental, podemos concluir que as variáveis Tipo de Palavra e Tamanho do Texto, colocadas em questão no presente artigo, têm grande influência no desempenho dos alunos.

O mesmo teste foi aplicado com os alunos do 1º ano do Ensino Médio e apresenta resultados diferentes em relação aos encontrados no grupo de alunos do Ensino Fundamental, sendo a condição TCF a que teve índices de erros mais elevados (Gráfico 2).

Diferentemente do que ocorreu com os alunos do Ensino Fundamental, as médias de erros em cada condição para os alunos do Ensino Médio foram muito próximas.

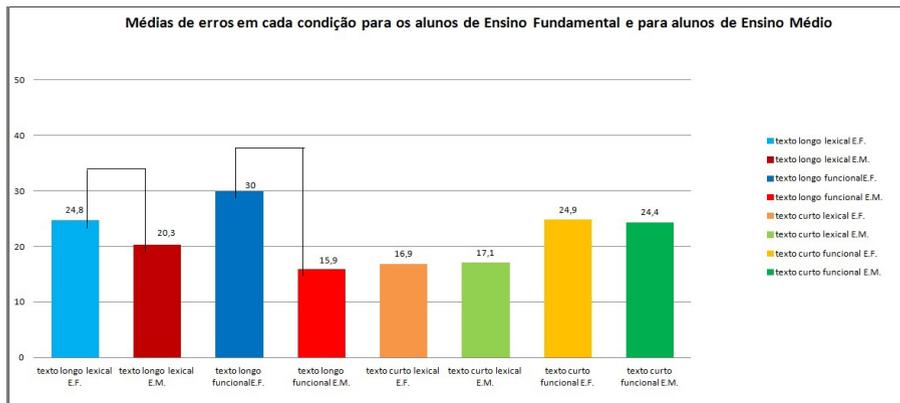
Em uma análise estatística, observamos que a comparação entre as condições não apresentou significância, pois as médias de erros foram muito próximas em todas as condições.

GRÁFICO 2 – Médias de erros no preenchimento das lacunas E. M.



Na análise estatística *Between Subjects* (comparação entre os dois grupos testados – alunos do Ensino Fundamental e alunos de Ensino Médio) fizemos, primeiramente, a análise estatística (T-teste). O T-teste mostrou que a comparação entre as condições com texto longo em que os alunos tinham que completar as lacunas com palavras funcionais foi significativa ($t(20)=4,43$ $p < 0,0003$), sendo os alunos do Ensino Fundamental os que apresentaram maiores médias de erro. Nos textos longos em que os alunos tinham de completar as lacunas com palavras lexicais, o resultado não foi significativo ($t(20)=1,93$ $p < 0,06$). No entanto, há fortes indícios de que há uma tendência, mostrando ser relevante a aplicação do teste com um número maior de alunos.

GRÁFICO 3 – Médias de erros no preenchimento das lacunas dos dois grupos



O teste de significância ANOVA (Gráfico 3) apresentou resultados significativos para a variável tipo de palavra $F(1,80)=6,31$ $p<0,01$, sendo as palavras funcionais as que tiveram maiores índices de erros, e para a variável nível de ensino $F(1,80) = 8,65$ $p<0,004$, sendo os alunos do ensino fundamental os que apresentam maiores índices de erros.

7 Discussão e agenda de pesquisa

De acordo com Kato ([1986] 2001), a compreensão da linguagem escrita envolve uma gama de diferentes tipos de conhecimentos. Em uma primeira análise, a linguagem escrita pode ser entendida como uma transcrição de uma língua. Nesse sentido, uma vez entendido o código, a proficiência em leitura deveria ser tão efetiva quanto a compreensão da língua. Por outro lado, a linguagem escrita abarca regras e especificidades que transcendem o que seria uma simples transcrição da língua, e demanda diferentes fatores cognitivos, como o acesso às memórias de curto e longo prazo e a capacidade de reconhecer e de estabelecer relações coesivas entre diferentes unidades significativas. Essas unidades tendem a variar em complexidade de acordo com o gênero textual e demandam maior grau de esforço formal quanto mais esses gêneros se distanciam da língua vernácula do leitor.

Sendo assim, medir a proficiência em leitura requer a capacidade de medir de modo controlado esses diferentes aspectos. O experimento

descrito na seção anterior mostrou que o teste de Cloze pode ser uma via para se manipularem diferentes aspectos relacionados à compreensão leitora. Nesse estudo, encontramos diferenças entre aspectos de ordem linguística (diferença entre palavras funcionais e lexicais) e aspectos de ordem não linguística (extensão do texto e escolaridade).

O desenho do experimento foi feito de modo a ser o mais fiel possível aos padrões de criação de testes de Cloze, utilizando o procedimento racional para simplesmente extrair palavras funcionais e lexicais, sem um olhar qualitativo sobre quais itens funcionais e quais itens lexicais seriam extraídos. Os resultados revelaram diferença quanto à variável tipo de palavra, o que tem o mérito de mostrar que o teste de Cloze pode ser sensível a essa variável de natureza linguística. Por outro lado, do ponto de vista do processamento da leitura, essa sensibilidade se torna mais relevante se pudermos avaliar o que essa diferença pode nos dizer sobre a compreensão leitora, como, por exemplo, determinar que tipo de cognição envolvida na leitura foi acessada para completar as lacunas.

Devido ao caráter multifacetado dos processos envolvidos na leitura, não é uma tarefa simples mensurar a relação entre testes de compreensão e os processos cognitivos que eles avaliam. Com relação especificamente aos processos cognitivos envolvidos na resolução do teste de Cloze, existe um debate quanto ao teste acessar primordialmente um conhecimento global contextual ou um processamento local imediato – que ativaria conhecimento prévio – na elaboração de respostas (SANTOS *et al.*, 2002). Conforme apontam Sousa e Hubner (2015, p. 37), “apesar de [o teste de Cloze] apresentar grande correlação com testes de compreensão leitora, há dúvidas sobre o que exatamente ele mede”.

A seguir, apresentamos uma breve análise qualitativa de alguns itens dos textos utilizados no experimento reportado, de modo a identificar a natureza da dificuldade dos participantes. O objetivo é mostrar que há especificidades em cada item que podem acessar aspectos diferentes do processamento da leitura. Analisamos qualitativamente, item a item, os erros que ocorrem tanto em palavras funcionais quanto em palavras lexicais. Esse tipo de análise revela especificidades de cada item, revelando que as habilidades requisitadas para completar uma lacuna podem ser de diferentes naturezas, ainda que se tratem de dois itens do mesmo tipo (lexical ou funcional). Esse tipo de análise é útil para informar

pesquisas futuras, de modo que se possam criar desenhos experimentais originais em que seja possível manipular especificamente o tipo de cognição envolvida, dando conta das possíveis variáveis intervenientes que estejam presentes em uma distinção de categorias como funcional e lexical. Os exemplos a seguir mostram alguns aspectos que podem ser explorados a partir dessas análises, e que podem ser manipulados de maneira controlada em testes futuros:

- 1) A lacuna em “sacar do _____” foi completada por participantes com os itens lexicais *cão, carro, bolso, cofre, homem, balcão, banco e caixa*.

No exemplo 1, o item no texto original é *revólver*, uma palavra lexical. Como se observa, o aluno não apresenta dificuldades em completar a lacuna quando se considera apenas o nível do sintagma, tendo êxito em reconhecer as restrições sintáticas e semânticas da estrutura, que licencia uma interpretação “tirar de x”. No entanto, na expressão idiomática “sacar do revólver”, trata-se de uma estrutura diferente, com a interpretação “pegar x”, em que revólver faz parte do objeto do verbo sacar e tem papel temático diferente de qualquer opção utilizada pelos alunos participantes do teste. Esse item é, portanto, um item que avalia uma erudição não esperada do falante comum e não revela exatamente um problema no processamento linguístico, já que os participantes tiveram êxito em completar o sintagma com itens possíveis. O uso desses itens possíveis revela, por outro lado, uma dificuldade do participante em termos de proficiência leitora. Ou seja, ainda que sejam todos itens possíveis para fechar o sintagma, e ainda que façam sentido dentro do campo semântico do texto, alguns desses itens lexicais não seriam aceitáveis levando-se em consideração a construção da coerência em uma análise mais macroestrutural.

Os participantes não parecem, portanto, ter problema em usar seu conhecimento linguístico para completar as lacunas do texto, o que revela que a dificuldade está relacionada a fatores extrínsecos à língua, associados possivelmente à estrutura textual ou ao desconhecimento de certas estruturas linguísticas estrangeiras à sua língua. Uma análise qualitativa permite, portanto, separar o tipo de erro que o aluno comete para que ações pedagógicas específicas sejam realizadas.

Uma análise preliminar de erros frequentes indica que, em geral, leitores em formação não apresentam problemas em completar sintagmas simples. Além disso, também não apresentam grandes dificuldades em identificar o tema predominante no texto, uma habilidade que depende apenas da ativação de um campo semântico por meio de um conjunto de itens lexicais, sem haver necessariamente uma estrutura coesiva entre eles. Por outro lado, é um desafio para leitores não experientes identificar unidades significativas maiores, como orações, sintagmas complexos – o que ocasiona dificuldade de compreensão geral do texto como uma unidade coesiva.

É interessante notar, ainda, em uma análise dos erros produzidos ao se completarem as lacunas com palavras funcionais, que a incidência de erro foi maior em estruturas de maior complexidade estrutural ou que dependem de maior vocabulário (exemplos 2 e 3 abaixo), exigindo maior nível de letramento – uma observação que é legitimada pela diferença encontrada entre os grupos de diferentes escolaridades.

- 2) Estruturas em que houve acerto: “um homem”; “o jardim”; “da casa”.
- 3) Estruturas em que houve erro: “ele astúcia” (em vez de “de astúcia”); “ele súplica” (em vez de “a súplica”); “disse que era da saúde pública ele indagou” (em vez de “disse que era da saúde pública e indagou”); “prometeu a nunca mais tentaria assaltar” (em vez de “prometeu que nunca mais tentaria assaltar”).

Para completar a lacuna em “___ astúcia”, o participante deve conhecer o item lexical “astúcia”, de frequência baixa nessa faixa etária. Ainda de menor frequência é a possibilidade de combinação com a preposição “de” para formar a expressão “de astúcia”. Além disso, é possível que, não reconhecendo a raiz do item lexical, o participante pode ter reconhecido a sequência final de letras da palavra como uma flexão verbal, o que justifica ter completado a lacuna com o item “ele”, que atuaria como sujeito desse suposto verbo. No caso de “súplica”, o participante pode ter falhado em reconhecer a diferença entre a forma verbal (não acentuada) e a forma nominal (acentuada) desse vocábulo. Novamente, se esses são os casos, o erro no preenchimento das lacunas decorre do não reconhecimento de um item lexical e não de uma falha no processamento linguístico. Em “prometeu a nunca mais tentaria assaltar”, o participante falhou em reconhecer a oração subordinada sendo usada

como argumento interno do verbo “prometeu”, e utilizou uma preposição (“a”) no lugar da conjunção integrante “que”. Nas estruturas presentes no exemplo 2, por outro lado, ainda que se trate de palavras funcionais, as estruturas são mais simples e as palavras de conteúdo são bastante conhecidas por estudantes do EF.

O desempenho melhor dos alunos do EM pode ter sido pelo tempo maior de exposição a estruturas mais complexas, incentivadas no processo de escolarização, e pelo aumento de vocabulário, que pode ser um fator relacionado tanto à idade quanto à escolaridade.

Ressaltamos, ainda, a relevância de associar o teste de Cloze com medidas cronométricas, que possibilitariam acessar de forma mais controlada a dificuldade do item. Por exemplo, acreditamos que um leitor experiente possivelmente também teria dificuldade em completar a lacuna apresentada no exemplo 1 com o item lexical “revólver”, mesmo que, havendo tempo para pensar, acabe acertando.

Em vista dessa breve análise de itens específicos, é possível constatar que o teste de Cloze pode ser uma ferramenta útil para avaliar diferentes aspectos da proficiência em leitura. Nossa breve análise indica que o teste, quando desenvolvido de forma consciente e utilizando-se o procedimento racional com resposta aberta, pode ser útil em revelar diferentes processos cognitivos, como, por exemplo, se o participante estaria acessando conhecimento global contextual ou conhecimento prévio (mais local, imediato). Isso só é possível, no entanto, se o teste for desenvolvido de maneira informada e qualitativa, controlando a criação das lacunas mais de modo a acessar o conhecimento que está sendo avaliado e menos em termos de categorias binárias, linguísticas ou não. Dessa forma, é possível identificar de maneira mais acurada se o erro ocorre por um problema em algum conhecimento linguístico específico, por um problema de proficiência em leitura, ou mesmo por uma lacuna no processo de letramento linguístico. A natureza de cada erro exige ações pedagógicas distintas.

8 Conclusão

Este trabalho se propôs discutir as possibilidades de se usar o teste de Cloze como ferramenta para medir aspectos relacionados à proficiência em leitura, verificando sua sensibilidade a aspectos de ordem linguística e não linguística e entretendo a possibilidade de se criarem

desenhos originais e qualitativos, de modo a determinar de forma mais granular os aspectos avaliados por cada item omitido.

Apresentamos uma revisão da literatura, em que procuramos descrever a técnica de Cloze em si, desde sua versão tradicional até versões com variação do procedimento. Apresentamos, também, estudos com o teste de Cloze em produções nacionais voltadas para o ensino fundamental e para o ensino médio.

A sensibilidade do teste foi medida por meio de um estudo que manipulou as variáveis tipo de palavra, tamanho do texto e escolaridade. O experimento envolveu dois grupos de participantes, um constituído por alunos do 6º ano do E.F. e o outro por alunos do 1º ano do E.M. Os resultados encontrados para o grupo do 6º ano (EF) confirmaram as hipóteses levantadas de que os participantes teriam mais dificuldade para compreender o Texto Longo por ser necessário se concentrar por um período de tempo maior e também apresentariam maior dificuldade para acessar as palavras funcionais, que podem ser facilmente justificadas devido à faixa etária compreendida por esses alunos e ao tempo de exposição a esses itens.

Em relação aos participantes do 1º ano (EM), a hipótese inicial era de que os participantes teriam mais facilidade no Texto Curto. No entanto, a média de erros dos resultados do Teste de Cloze foi maior no Texto Curto (24,4) em comparação com o Texto Longo (20,3), mesmo não sendo uma diferença significativa, foi um resultado inesperado. Uma possível justificativa é que o Texto Longo tenha proporcionado mais informações ao participante que teve um embasamento maior para acessar e preencher as lacunas do teste com mais acurácia.

Quanto à comparação entre os grupos, o melhor resultado foi encontrado no grupo dos participantes do Ensino Médio. Fato que já era esperado devido não só ao espaço temporal que separa os dois grupos, mas também à quantidade de anos de escolaridade, visto que do 6º ano (EF) para o 1º ano (EM) são três anos a mais em contato e interação com o ambiente acadêmico.

Ao final, destacamos a importância de se desenvolverem testes de Cloze que possibilitem acessar aspectos específicos relacionados à proficiência leitora, de modo a identificar o que está sendo avaliado pelo teste. Dessa forma, é possível identificar e trabalhar a natureza do problema em leitura apresentado pelo participante. Assim, o teste de Cloze pode se consolidar como sendo uma ferramenta importante e

versátil, capaz de acessar a proficiência leitora em níveis mais locais, o que pode revelar problemas de processamento linguístico ou simples desconhecimento de itens lexicais, ou em níveis mais macroestruturais, que podem revelar lacunas no letramento em leitura.

Agradecimentos

As autoras agradecem a equipe do CIEP 246 – Professora Adalgisa Cabral de Faria pela permissão do desenvolvimento da pesquisa. Agradecem também a colega Prof. Dr. Marília Costa (UFRJ) que destinou tempo à leitura da versão anterior e aos pareceristas anônimos por suas variadas sugestões.

Referências

ABRAHAM, R. G.; CHAPELLE, C. A. The meaning of Cloze test scores: An item difficulty perspective. *The Modern Language Journal*, Wiley Online Library, v. 76, n. 4, p. 468-479, 1992.

ADELBERG, A. H.; RAZEK, J. R. The Cloze Procedure: A Methodology for Determining the Understandability of Accounting Textbooks. *The Accounting Review*, American Accounting Association, v. 59, n. 1, p. 109-122, jan. 1984.

BALIZA, A. A.; SILVA, D. V. Avaliação da Compreensão em Leitura em Estudantes do Ensino Fundamental. *Educare, Revista Científica de Educação*, UNASP, v.1, n.1, p. 93-114, 2015. ISSN 2447-5432.

BORMUTH, R. J. Cloze test readability: criterion reference scores. *Journal of Educational Measurement*, Wiley Online Library, v. 5, n. 3, p. 189-196, 1968.

BRAINE, M.D.S. What is learned in acquiring word classes: A step toward an acquisition theory. In: MacWHINNEY, B. (Ed.). *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum, 1987. p. 65-87.

BRUSINI, P.; DEHAENE-LAMBERTZ, G.; VAN HEUGTEN, M.; DE CARVALHO, A.; GOFFINET, F.; FIVET, A. C.; CHRISTOPHE, A. Ambiguous function words do not prevent 18-month-olds from building accurate syntactic category expectations: an ERP study. *Neuropsychologia*, Elsevier, v. 98, p. 4-12, April 2017. <http://dx.doi.org/10.1016/j.neuropsychologia.2016.08.015>.

BRUSINI, P.; DEHAENE-LAMBERTZ, G.; DUTAT, M.; GOFFINET, F.; CHRISTOPHE, A. ERP evidence for on-line syntactic computations in 2-year-olds. *Developmental Cognitive Neuroscience*, Elsevier, v. 19, p. 164-173, 2016. <https://doi.org/10.1016/j.dcn.2016.02.009>.

CAIN, K.; NASH, H. The influence of connectives on young readers' processing and comprehension of text. *Journal of Educational Psychology*, American Psychological Association, v. 103, p. 429-441, 2011. doi: <https://doi.org.10.1037/a0022824>.

CEREJA, W. R.; MAGALHÃES, T.C. *Português: linguagens*. 6º ano. 5. ed. reform. São Paulo: Atual. 2009.

COELHO, N. N. *Literatura Infantil: teoria, análise, didática*. São Paulo: Moderna, 2000.

COHEN, J. H. The effect of content area material on Cloze test performance. *Journal of Reading*, International Literacy Association, v. 19, n. 3, p. 247-250, 1975.

COLLINS. *Escolar Plus Dictionary: English/Portuguese*. [s.l.]: Heinle Cengage Learning, 2009.

CONDEMARÍN, M.; MILICIC, N. *Teste de Cloze: aplicações psicopedagógicas*. Santiago de Chile: Editorial Andrés Bello, 1988

CUNHA, N. de B. Diagnóstico de compreensão de leitura por meio do teste de *Cloze*. In: CONGRESSO DE LEITURA DO BRASIL, 17., 2009, Campinas. *Anais...* Campinas: Unicamp, 2009. v.1, p.238-247.

CYRANKA, L. F.; PERNAMBUCO, D. L. C. A língua culta na escola: uma interpretação sociolinguística. *Instrumento: Revista de Estudo e Pesquisa em Educação*, Juiz de Fora, v. 10, p.17-28, jan./dez. 2008.

DIAS, N. M.; LEÓN, C. B. R.; PAZETO, T. de C. B.; MARTINS, G. L. L.; PEREIRA, A. P. P.; SEABRA, A. Avaliação da leitura no Brasil: revisão da literatura no recorte 2009-2013. *Revista Psicologia: Teoria e prática*, São Paulo, v. 18, n. 1, p. 113-128, jan./abr. 2016.

FEDERMEIER, K. D.; WLOTKO, E. W.; OCHOA-DEWALD, E.D.; KUTAS, M. Multiple effects of sentential constraint on Word processing. *Brain Research*, Elsevier, v. 1146, p.75-84, 2007. <https://doi.org/10.1016/j.brainres.2006.06.101>.

FONSECA, R. P.; CASARIN, F. S.; OLIVEIRA, C. R.; GINDRI, G.; SOARES, E. C. S.; ORTIZ, K. Z.; PARENTE, M. A. M. P.; SCHERER, L. C. Adaptação de instrumentos neuropsicológicos verbais: um fluxograma de procedimentos para além da tradução. *Interação em Psicologia*, Universidade Federal do Paraná, v. 15, n. especial, p. 59-69, 2011. <https://doi.org/10.5380/psi.v15i0.25374>.

GONZALEZ, C. A.; QUADROS, E. S.; SCHWINDT, L. C. S.; TOLEDO, E. E. A influência da variável escolaridade em fenômenos fonológicos variáveis: efeitos retroalimentadores da escrita. *Revista Virtual de Estudos da Linguagem – Revel*, [s.l.], v. 5, n. 9, ago. 2007.

INFANTE, U. *Textos: leituras e escritas: literatura, língua e redação*. São Paulo: Scipione, 2000. v. 1.

JOLY, M. C. R. A. *Sistema Orientado do Cloze*. Relatório de pesquisa apresentado à Universidade São Francisco, Itatiba, São Paulo, 2006.

JOLY, M. C. R. A. The validity of Cloze Oriented System (COS): a correlation study with an electronic comprehension test and a reading attitude survey. *Psicologia Escolar e Educacional*, Campinas, v. 11, n. especial, p. 49-58, 2007.

JOLY, M. C. R. A. *Programa Informatizado de Leitura Estratégica (PILE) – Pesquisa em desenvolvimento*. Itatiba, SP: Universidade São Francisco, 2008.

JOLY, M. C. R. A. Estudos com o Sistema Orientado de Cloze para o ensino Fundamental. In: SANTOS, A. A. A.; BUCHOVITCH, E.; OLIVEIRA, K. L. (Ed.). *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. p. 119-144.

JOLY, M. C. R. A.; PIOVEZAN, N. M. Avaliação do Programa Informatizado de Leitura Estratégica para estudantes do ensino fundamental. *Paidéia*, Ribeirão Preto, v. 22, n. 51, p. 83-90, abr. 2012. <https://doi.org/10.1590/S0103-863X2012000100010>.

JOLY, M. *et al.* Avaliação da Compreensão de Leitura pelo Sistema Orientado de Cloze (SOC). *Fractal: Revista de Psicologia*, Universidade Federal Fluminense, v. 26, n. 1, p. 223-242, jan./abr. 2014.

KATO, M. A. *No mundo da escrita: uma perspectiva psicolinguística*. 7. ed., 3ª impressão. São Paulo: Ática, [1986], 2001.

KINTSCH, W.; VAN DIJK, F. A. Toward a model of text comprehension and production. *Psychological Review*, American Psychological Association, v. 85, p. 363-339, 1978.

KLEIMAN, A. *Texto e leitor: aspectos cognitivos da leitura*. Campinas: Pontes, 1989.

OLIVEIRA, K. L.; BORUCHOVITCH, E.; SANTOS, A. A. A. A técnica de Cloze na avaliação da compreensão em leitura. In: SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. p. 47-77.

OLIVEIRA, K. L.; CANTALICE, L. M.; FREITAS, F.A. Compreensão em leitura no ensino médio: análise de acertos por item. In: SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. p. 165-185.

PAGE, W. D. The post-oral reading Cloze test: New link between oral reading and comprehension. *Journal of Reading Behavior*, Chicago, v. 7, n. 4, p. 383-389, 1975. <https://doi.org/10.1080/10862967509547155>.

PARENTE, M. A. M. P.; SCHERER, L. C.; ZIMMERMANN, N.; FONSECA, R. P. Evidências do papel da escolaridade na organização cerebral. *Revista Neuropsicologia Latinoamericana* [online], Sociedad Latinoamericana de Neuropsicologia, v. 1, n. 1, p. 72-80, 2009. Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S2075-94792009000100009&lng=pt&nrm=isso>. Acesso em: 3 jan. 2017. ISSN 2075-9479.

PEREIRA, M.; ARAÚJO, A. Considerações acerca da variável escolaridade e sua influência sobre a variação entre verbo-sujeito na 3ª pessoa do plural no português brasileiro. *PERcursos Linguísticos*, Vitória - ES, v. 6, n. 12, 2016. ISSN: 2236-2592.

SADEGHI, K. Phrase Cloze: a better measure of reading? *The Reading Matrix*, Reading Matrix Inc., v. 14, n. 1, p. 76-94, April 2014.

SANTOS, A. A. A. *Diagnóstico de compreensão de leitura por meio do teste de Cloze*. Programa de Pós Graduação Stricto Sensu em Psicologia. Itatiba: Universidade São Francisco, 2002.

SANTOS, A. A. A. O Cloze como técnica de diagnóstico e remediação da compreensão em leitura. *Interação em Psicologia*, UFPR, v. 8, n. 2, p. 217-226, jul./dez. 2004. <https://doi.org/10.5380/psi.v8i2.3257>.

SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. Análise da fidedignidade entre dois tipos de pontuação do teste de Cloze. *Psicologia em Pesquisa*, Universidade Federal de Juiz de Fora, v. 1, n. 1, p. 41-51, jan./jun. 2007.

SANTOS, A. A. A.; BORUCHOVITCH, E.; OLIVEIRA, K. L. *Cloze: um instrumento de diagnóstico e intervenção*. São Paulo: Casa do Psicólogo, 2009. 380 p.

SANTOS, A. A. A.; PRIMI, R.; TAXA, F. O. S.; VENDRAMINI, C. M. M. O teste de Cloze na avaliação da compreensão em leitura. *Psicologia: Reflexão e Crítica*, Universidade Federal do Rio Grande do SUL, v. 15, n. 3, p. 549-560, 2002. <https://doi.org/10.1590/S0102-79722002000300009>.

SEGALOWITZ; S. J.; LANE, K. Lexical access of function versus content words. *Brain and Language*, Elsevier, v. 75, n. 3, p. 376-389, Dec. 2000.

SMITH, F. *Compreendendo a leitura: uma análise psicolinguística da leitura e do aprender a ler*. Tradução de Daise Batista. 4. ed. Porto Alegre: Artmed, 2003.

SOUSA, L. B. de; HUBNER, L. C. Desafios na avaliação da compreensão leitora: demanda cognitiva e leitorabilidade textual. *Revista Neuropsicologia Latinoamericana*, Sociedad Latinoamericana de Neuropsicologia, v. 7, n. 1, p. 34-46, 2015.

SUEHIRO, A. C. B. Produção Científica sobre o Teste de Cloze. *Psicologia Escolar Educacional*, Maringá, v. 17, n. 2, p. 223-232, dez. 2013. <https://doi.org/10.1590/S1413-85572013000200004>.

SUEHIRO, A. C. B.; BORUCHOVITCH, E. Compreensão em Leitura em Estudantes do Terceiro e Quarto Ciclos do Ensino Fundamental. *Psico-USF*, Itatiba, v. 21, n. 3, p. 561-572, dez. 2016.

SWAFFAR, J. K.; ARENS, K. M.; BYRNES, H. *Reading for meaning: an integrated approach to language learning*. Englewood Cliffs, N.J.: Prentice Hall, 1991.

TAFT, M. Lexical processing of functionally constrained words. *Journal of Memory and Language*, Elsevier, v. 29, p. 245-257, 1990. [https://doi.org/10.1016/0749-596X\(90\)90075-B](https://doi.org/10.1016/0749-596X(90)90075-B).

TAYLOR, W. L. Cloze procedure: a new tool for measuring readability. *Journalism Quarterly*, Questia Trusted Online Research, v. 30, p. 415-433, 1953.

VAN HEUGTEN, M.; CHRISTOPHE, A. Infants' acquisition of grammatical gender co-occurrences. *Infancy*, Wiley Online Library, v. 20, p. 675-683, 2015.

VAN PATTEN, B. *Input processing and grammar instruction: theory and research*. Norwood, NJ: Ablex, 1996.

VOTRE, S. J. Relevância da variável escolaridade. In: MOLLICA, M. C.; BRAGA, M. L. (Org.). *Introdução à Sociolinguística: o tratamento da variação*. São Paulo-SP. Editora Contexto, 2012. p. 51-57.

O movimento ocular na leitura realizada por revisores de textos

Revisers' eye movement in reading

Délia Ribeiro Leite

Conselho Regional de Contabilidade de Minas Gerais, Belo Horizonte, Minas Gerais
/ Brasil

dribeiroleite@yahoo.com.br

José Olímpio de Magalhães

Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais / Brasil

joseolimpiomagalhaes@yahoo.com.br

Resumo: Este estudo objetivou investigar o processamento da leitura de profissionais revisores de textos, comparando o movimento ocular destes com o de sujeitos que não trabalham profissionalmente com revisão (não revisores). Na tarefa de leitura e detecção de erros realizada por 14 revisores e 14 não revisores, os participantes deveriam ler e revisar pequenos textos jornalísticos projetados no computador, clicando com o mouse nos erros e/ou inadequações que encontrassem. Havia dois tipos de erros nos estímulos experimentais: a) supressão de preposição e b) incoerência gerada por uma anáfora nominal incorreta. Na análise estatística, quando garantida a normalidade, foram realizados modelos mistos; do contrário, a análise foi realizada por meio de testes não paramétricos. Quanto à análise do movimento ocular, a perspectiva geral foi de que os revisores apresentaram valores mais elevados, o que significa que eles foram mais lentos na leitura do que os não revisores, o que foi identificado nas medidas do nível do texto, da sentença e do trecho alvo. Esta pesquisa contribui para uma caracterização dos processos de leitura envolvidos em tarefas de revisão de textos por revisores profissionais.

Palavras-chave: revisão de textos; revisão profissional; movimento ocular; proficiência.

Abstract: This study aims to investigate the reading processing of professional revisers comparing their eye movement with subjects who do not work professionally in revising (non revisers). In the reading and error detection task proposed and performed by 14 revisers and 14 non revisers, the participants had to read and edit short journalistic texts displayed on a computer, by clicking on the errors and/or inadequacies they found using the mouse. There were two types of errors: a) suppression of preposition and b) incoherence generated by an incorrect nominal anaphora. In the statistic analysis, when normality was ensured, mixed models were carried out; otherwise, the analysis was carried out by means of non-parametric tests. Considering the eye movement analysis, the general perspective indicate that revisers presented the highest values, meaning they were slower in reading than non revisers, which occurred concerning the measures of the text, sentence and local levels. This research contributes to a characterization of reading processes involved in revisions of texts by professional reviewers.

Keywords: revision; professional revisers; eye movement; proficiency.

Recebido em: 9 de dezembro de 2016.

Aprovado em: 9 de janeiro de 2017.

1. Introdução

A revisão está diretamente relacionada com o processo de leitura e produção de textos. Por isso, ao relermos um texto que produzimos, buscamos assumir uma visão exotópica¹ e verificar aspectos diversos da estrutura textual, de forma a identificar possíveis erros, bem como trechos cuja interpretação seja difícil para o leitor. No entanto, nem sempre o autor é o melhor leitor para revisar o seu próprio texto, seja por questões técnicas, relativas à falta de conhecimento específico no campo dos estudos da linguagem (no caso de autores de outras áreas), seja por questões práticas: a exotopia pode ficar comprometida, tendo em vista que o texto, ao ser lido pelo próprio autor, perde a imprevisibilidade, o que pode fazer com que as inadequações de linguagem não sejam percebidas.

¹ A visão exotópica é aquela em que a leitura é realizada com um olhar diferente daquele do produtor do texto, um olhar que simularia a leitura realizada pelo leitor para o qual aquele texto é destinado.

Por isso, há o profissional revisor de textos, cuja ocupação é a leitura dos textos produzidos por outrem, de forma a identificar inadequações de linguagem e propor correções ou alterações diversas.

Apesar de ser uma atividade antiga e importante, poucos são os estudos cujo objeto é a revisão profissional, muito embora haja um imaginário de que essa tarefa exige uma demanda e uma qualificação especiais, principalmente no que se refere à leitura: “O revisor não lê como os demais homens leem, ele fotografa a palavra visualmente; e o texto reflete-se em sua correção.” (WAGNER; CUNHA, 2012, p. 12).

Dessa forma, este trabalho propõe-se a investigar o processamento da leitura realizada por profissionais revisores de textos. Para tanto, foi utilizado como método de investigação o rastreamento ocular. Verificamos os padrões de fixações e sacadas realizados por esses profissionais durante a revisão, bem como se esses padrões se aproximaram do depreendido quando indivíduos que não trabalham profissionalmente com revisão leem os mesmos textos. Portanto, este trabalho tem como objetivos:

- Investigar os padrões do movimento dos olhos na leitura realizada por profissionais revisores de textos, em uma tarefa de detecção de erros;
- Comparar o movimento ocular de profissionais revisores de textos com o de sujeitos que não realizam profissionalmente a revisão, em uma tarefa de leitura e detecção de erros;
- Verificar se os profissionais revisores de textos e se os sujeitos que não realizam a revisão profissional fazem uma leitura voltada tanto para a superfície textual quanto para níveis mais globais do texto.

A hipótese assumida neste estudo é a de que a leitura realizada pelos revisores de textos é menos automática, mais controlada e mais detalhada, o que acarreta mais proficiência na atividade de revisão.

2. A revisão de textos sob uma abordagem cognitiva

Heurley (2006) destaca três principais visões da revisão na perspectiva cognitiva: a revisão como uma modificação efetiva de um texto; a revisão como subprocesso do processo de escrita, visando melhorar o texto já escrito; e a revisão como componente distinto da

produção escrita, que envolveria um conjunto de subprocessos e de procedimentos implicados no controle da escrita.

Hayes *et al.* (1987), por sua vez, propõem um modelo especificamente de revisão. O modelo é dividido em processos e conhecimentos. O primeiro processo envolvido é a definição da tarefa, que abarca questões como os objetivos do revisor, os traços do texto a serem examinados e a forma como se pretende proceder à revisão. O segundo é a avaliação, quando são selecionados os objetivos da leitura: compreensão, avaliação ou detecção de problemas. Com base na representação que se faz do problema, será selecionada uma estratégia, sendo possível: modificar ou controlar o processo de revisão em si ou modificar o texto. Ao modificar o processo de revisão, o revisor pode ignorar o problema, buscar mais informações de forma a melhorar o seu diagnóstico ou adiar a ação, inclusive quando opta por fazer mais de uma leitura, uma delas direcionada a um nível mais alto e outra relacionada com os aspectos da superfície textual. Ao contrário, se opta por modificar o texto, o revisor pode reescrevê-lo ou revisá-lo. Nesse caso, a revisão comporta os casos em que o revisor corrige as inadequações mantendo o máximo possível do texto original. Uma questão importante é que os revisores podem redefinir a tarefa ao longo da realização da revisão e, assim, o conhecimento utilizado, que inclui objetivos, critérios e restrições, é modificado dinamicamente durante a revisão. A revisão está relacionada com a leitura para detecção de problemas, que, de acordo com os autores, é diferente, por exemplo, daquela direcionada somente à compreensão.

Em 1996, Hayes propõe uma reformulação do modelo de escrita de Hayes e Flower (1980), com o objetivo principal de acrescentar elementos cognitivos importantes, como a memória de trabalho, assim como reorganizar os elementos existentes de forma a mostrar a relação deles com processos cognitivos mais gerais envolvidos em vários tipos de atividades. A revisão passa a ser vista como uma forma de interpretação de texto. O autor propõe um novo modelo de revisão, no qual há uma estrutura de controle, composta pelo esquema da tarefa de revisão; processos fundamentais, entre eles a reflexão, o processamento e a produção do texto; e os recursos utilizados, que são a memória de trabalho e a memória de longa duração. A revisão integra juntamente com a reflexão e a produção do texto o subcomponente dos processos cognitivos, o qual, por sua vez, está inserido em um componente maior.

Este, além dos processos cognitivos, envolve um subcomponente afetivo/motivacional, a memória de trabalho e a memória de longo prazo. Assim, o autor considera que, para entender a revisão, é necessário entender não só o processo, mas também sua estrutura de controle e os recursos envolvidos. O autor pontua que esse conjunto é adquirido com a prática, o que se depreende, inclusive, da diferença entre revisores e produtores de textos experientes e novatos explorada em Hayes *et al.* (1987).

Como se vê, no âmbito desses estudos, a revisão deixou de ter um papel secundário no processo de escrita e passou a ser vista como uma etapa importante desse processo, que atua em vários níveis. Cada vez mais, ela tem sido tratada como uma atividade de controle, que pode operar separadamente ou ao longo do processo de escrita.

3. Movimento ocular e leitura

Os olhos se movimentam em sacadas intercaladas com fixações. Enquanto a sacada corresponde a um rápido movimento dos olhos para mover o foco de uma área para outra, a fixação equivale ao tempo gasto focalizando-se uma determinada área. É possível a verificação de variáveis temporais e espaciais quando se registra o movimento dos olhos, já que há um deslocamento espacial, ou seja, para onde será direcionado o movimento, e um deslocamento temporal, representado pelo momento em que ocorrerá esse movimento. Além disso, sabe-se que o tipo de movimento depende bastante da informação requerida (LAND, 2007, p. 78).

Além disso, nem sempre o movimento ocular ocorre de maneira linear, pois há situações em que são realizadas regressões para locais que já foram focalizados. De acordo com Luegi (2006, p. 23), “Cerca de 15% dos movimentos sacádicos durante a leitura são movimentos regressivos, ou seja, partem da direita para a esquerda, para regiões anteriores do texto, na mesma linha ou em algumas linhas acima”.

As regressões podem ser correções de sacadas que foram muito longas e, em função disso, é necessário que os olhos façam um breve retorno. Podem, também, ser decorrentes de dificuldades de entendimento de alguma parte do texto. Nesse caso, geralmente, as regressões são longas (maiores que 10 caracteres na mesma linha ou em direção a linhas anteriores) e o comportamento com relação ao retorno varia conforme a proficiência do leitor.

O registro do movimento dos olhos é utilizado em uma grande gama de estudos sobre o processamento da linguagem. Muitas variáveis podem influenciar os valores das fixações e das sacadas, inclusive se se trata de uma leitura em voz alta ou silenciosa. Como não há estudos específicos sobre o movimento ocular durante a revisão de textos, nesta seção, serão apresentados estudos que abordam questões relacionadas aos tipos de erros presentes nos textos experimentais deste trabalho, bem como fatores considerados na análise depreendida neste trabalho, como os retornos e a releitura dos textos.

Como o movimento dos olhos se processa em fixações e sacadas, de uma fixação para outra podem ser puladas palavras. Sabe-se que palavras funcionais tendem a ser puladas, ou seja, não serem focalizadas. De acordo com Rayner, as palavras funcionais são fixadas aproximadamente 35% das vezes (RAYNER, 1998, p. 375). Isso pode ser ocasionado por elas serem mais facilmente identificáveis, ou seja, mais previsíveis pelo contexto, ou mesmo pelo fato de elas serem mais frequentes (STAUB; RAYNER, 2007). O fato de elas serem curtas também pode influenciar, já que, à medida que o tamanho da palavra aumenta, a probabilidade de ela ser fixada também aumenta – palavras de 2 a 3 letras são fixadas, em geral, 25% das vezes (RAYNER, 1998, p. 375).

Hyönä e Nurminen (2006) investigam as regressões e, em seus resultados, os leitores que tendem a direcionar as regressões para partes realmente informativas do texto conseguem apresentar um resumo melhor daquilo que leram. Dessa forma, há também aspectos relacionados a níveis textuais mais altos que interferem no movimento dos olhos. Os autores pesquisam como leem leitores competentes, com o objetivo de confirmar a classificação proposta em estudo anterior (HYÖNÄ; KAAKINEN; LORCH, 2002) e verificar se os próprios leitores têm consciência das estratégias de leitura que eles utilizam. Foram identificados três padrões de leitores: os processadores de estruturas de tópico, que foram minoria, seguidos dos leitores lineares rápidos e, por fim, dos leitores lineares lentos, que foram maioria. Os leitores têm consciência do tempo demandado na leitura, se são ou não lentos, bem como sabem se fazem ou não retornos a partes do texto, porém não têm uma noção precisa de quais os locais para os quais eles fazem retornos. Portanto, há diferenças idiossincráticas relacionadas ao movimento ocular e à leitura, embora o direcionamento dado na tarefa possa influenciar os padrões. No caso da

pesquisa acima, os leitores tiveram que fazer um resumo do texto, o que pode ter influenciado a maneira como eles se comportaram.

Hyönä e Niemi (1990) investigam a releitura de um texto. A releitura faz com que o processo seja facilitado e, com isso, todas as medidas investigadas são diminuídas, tanto o número e o tempo de fixações quanto o número de regressões. Os autores também investigam se a releitura influencia na tendência em se fixarem mais os trechos com informações importantes. Nesse sentido, a conclusão é de que as partes mais importantes tendem a ser mais fixadas, mas isso ocorre tanto na primeira leitura quanto na releitura. No entanto, a releitura facilita o processamento das sentenças que são mais informativas do que daquelas não informativas. Além disso, para cada leitura sucessiva, a média das fixações é mais longa no início dos textos do que no final deles.

Vauras, Hyönä e Niemi (1992) investigam a leitura de textos coerentes e incoerentes. Nos textos incoerentes, há a mudança da ordem das sentenças. Os autores utilizam como metodologia o registro do movimento dos olhos e também uma medida off-line, que é a análise de reescritas dos textos realizadas pelos leitores. Eles concluem que incoerências aumentam tanto o tempo de fixação quanto as regressões. No entanto, embora em textos coerentes o aumento do tempo de fixação ocasione melhora na reescrita da parte fixada, o mesmo não ocorre com textos incoerentes.

Luegi (2006) pesquisa o registro do movimento dos olhos entre leitores do português europeu. A autora apresenta sentenças modificadas com a finalidade de se criarem inconsistências (agramaticalidade ou ambiguidade) e verifica a reação dos leitores. Além disso, essas manipulações são realizadas em dois textos: um em que predominam termos técnicos e outro cujo tema é mais comum. A autora verificou que o efeito das variáveis só é visível quando se trata do texto com vocabulário técnico e manipulação das sentenças gerando inconsistências. Esses resultados são vistos tanto no número de fixações quanto no tempo total de leitura. Além disso, a autora verifica que os leitores são bastante precisos no que se refere às regressões e retornam especificamente para o local onde há problemas.

Alamargot *et al.* (2006) registram concomitantemente o movimento ocular e a execução grafomotora para investigar o processo de escrita. Embora o foco do trabalho não seja a revisão de textos, eles observam que a detecção de erros tipográficos pode ocorrer em tempos

muito curtos, como o de uma fixação na leitura. Pode, ainda, ocorrer em paralelo com a execução grafomotora.

Portanto, o registro do movimento dos olhos pode indicar aspectos importantes do processamento da leitura. Dessa forma, os trabalhos experimentais demonstram que tanto a fisiologia da visão quanto o processamento da leitura são determinantes quanto ao padrão do movimento dos olhos adotado pelos sujeitos. Esses estudos dão suporte à hipótese deste trabalho de que haverá um padrão de leitura diferenciado conforme o tipo de erro e, também, em função da experiência dos sujeitos em revisão de textos.

4. Metodologia

4.1 Participantes

Os sujeitos participantes da pesquisa foram divididos em dois grupos: não revisores e revisores, sendo que cada grupo foi composto por 14 sujeitos. No grupo dos não revisores, participaram da pesquisa estudantes de ensino superior que não cursavam Letras nem Comunicação Social, pois a formação nesses dois cursos é pré-requisito para a investidura de vários cargos de revisão e, com isso, esse estudante tem uma formação mais aprofundada em estudos de linguagem, o que o diferencia de um grupo que representa sujeitos não revisores. Para a escolha dos revisores, foi pré-requisito que o sujeito trabalhasse profissionalmente com revisão, em órgãos públicos, na iniciativa privada ou como autônomo. Todos os revisores tinham formação em Letras, sendo que um deles estava no último período do curso, mas já acumulava 3 anos de experiência em revisão de textos. Com relação ao tempo de experiência, 5 deles trabalhavam com revisão de 1 a 4 anos; 5, de 5 a 10 anos; e 4, por mais de 10 anos.

4.2 Construção dos estímulos

Os textos, do gênero jornalístico, redigidos especialmente para o experimento, consistiram de um parágrafo informativo. Os problemas dos textos experimentais foram de dois tipos:

- supressão de preposição;
- incoerência gerada por uma anáfora nominal incorreta.

Tais tipos de problemas foram selecionados em função do custo do processamento. Esses problemas representam dois níveis diferentes: a supressão de preposição relaciona-se ao nível da palavra e do sintagma; e a incoerência gerada por uma anáfora nominal incorreta, ao nível do texto, pois é necessário que seja realizada a integração entre as sentenças do texto.

Além disso, foram inseridos textos sem erros e com outros tipos de erros, considerados distratores. Em cada sessão experimental, foram apresentados 20 textos experimentais, 10 com supressão de preposição e 10 com anáfora incorreta, nos quais não havia outros erros/inadequações previstos de serem detectados. Além disso, havia mais 30 textos distratores (8 sem erros, 2 com um erro, 7 com dois erros, 9 com três erros e 4 com quatro erros).²

Nos exemplos 1 e 2, abaixo, são apresentados dois exemplos de textos experimentais, um com supressão de preposição (presença duas torcidas) e outro com anáfora incorreta (americano retomando argentino).

Exemplo 1 – Texto experimental – supressão de preposição

Após um longo período, o clássico entre Atlético Mineiro e Cruzeiro contará com a presença duas torcidas mineiras. O Cruzeiro será o mandante da partida e vai disponibilizar uma parte dos ingressos para a torcida adversária. O jogo do Campeonato Mineiro será na próxima semana, na inauguração do novo Mineirão, e terá policiamento reforçado.

Exemplo 2 – Texto experimental – anáfora incorreta

Na noite de ontem, um bebê foi encontrado próximo às margens da Lagoa da Pampulha por um argentino. O bebê estava enrolado em uma manta, e a mãe da criança foi presa horas depois do ocorrido. O americano confirmou em depoimento que a mãe abandonou a criança no local e fugiu logo em seguida.

Os textos experimentais e os distratores foram distribuídos entre 8 temas. Na construção dos textos experimentais, foram controladas variáveis que podem influenciar os resultados: 3 sentenças; 54 palavras

² Não foi feita uma distribuição em Quadrado Latino, com a leitura dos mesmos textos com e sem erros, porque isso exigiria um número maior de participantes ou mesmo um número maior de textos experimentais a serem lidos por sessão, o que seria inviável, tendo em vista a dificuldade na participação de revisores profissionais, os quais tinham que se deslocar até o local do experimento, e o cansaço gerado em função de um experimento muito longo.

por texto, 18 por sentença; a posição das palavras alvo, que foram a 15^a e a 16^a do texto na supressão de preposição e, nos textos com anáfora incorreta, a 38^a, sendo que o referente incorreto era a 18^a palavra do texto. Na supressão de preposição, a palavra anterior à preposição tinha 3 sílabas e a posterior, 2. Já nos textos em que havia a anáfora incorreta, tanto o referente quanto a anáfora tinham 4 sílabas.

Especificamente com relação aos textos com anáfora incorreta, levou-se em consideração, também, a estrutura informacional dos textos: o referente (item lexical novo introduzido no texto na primeira sentença) não é o sujeito da oração e vem acompanhado por um artigo indefinido. Portanto, não está focalizado. Já a anáfora incorreta é o sujeito da terceira oração e vem acompanhada por um artigo definido, com o objetivo de que se tratasse de uma retomada de um item anteriormente introduzido no discurso. Cuidou-se de compor textos distratores contendo sentenças com uma estrutura sintática semelhante, mas sem uma anáfora incorreta, de forma a minimizar as possibilidades de os participantes preverem a existência do erro em função da estrutura sintática.

Como a frequência de uso está bastante relacionada ao tempo de fixação, foi utilizado o corpus Banco de Português (BP), compilado por Tony Sardinha da PUC-SP, para se controlar essa variável. Em geral, as palavras experimentais na supressão de preposição são mais frequentes do que na anáfora incorreta.

Foi utilizado, ainda, o programa livre “Coh-Metrix-Port”,³ o qual analisa vários parâmetros, visando, principalmente, oferecer subsídios para a análise da legibilidade dos textos. O índice Flesch é uma medida superficial da inteligibilidade de um texto, que leva em consideração os tamanhos médios de palavras e sentenças, correlacionando-os com a facilidade de leitura. Quanto mais alto o valor obtido, mais fácil o processamento do texto. Considerando-se esse índice, os textos do experimento são classificados como muito fáceis, fáceis ou difíceis, o que correspondente às séries anteriores ao ensino superior. Como os sujeitos participantes da pesquisa são, no mínimo, universitários, não deveriam ter dificuldade para ler o texto, levando-se em conta esse índice.

³ Disponível em <www.nilc.icmc.usp.br/coh-metrix-port>. A ferramenta Coh-Metrix foi desenvolvida na Universidade de Memphis e calcula a coerência de textos, usando diversas medidas. No Brasil, a adaptação para o português foi realizada por pesquisadores do Núcleo Interinstitucional de Linguística da Universidade de São Paulo – campus de São Carlos.

Como a previsibilidade das palavras pode influenciar o tempo de fixação nelas, foi controlado se a palavra alvo dos textos experimentais era previsível, por meio de um teste de completção (*frame*), realizado com alunos do curso de Letras da UFMG. Para a supressão de preposição, a palavra alvo, ou seja, aquela posterior à preposição, é previsível, pois a palavra correta foi a preenchida entre 53 a 100% das vezes. Já na anáfora incorreta, as palavras alvo (o referente e a anáfora) não são previsíveis, já que as escolhidas foram indicadas de 0 a 28% das vezes.

Por fim, antes do início da coleta de dados, foi realizado, um pré-teste com 16 alunos do curso de pós-graduação em Revisão de Textos da PUC-Minas, visando verificar quais erros/inadequações seriam detectados tanto nos textos experimentais quanto nos distratores. Com isso, foi possível realizar as últimas adequações nos textos experimentais, de forma a minimizar as possibilidades de os sujeitos fazerem detecções inesperadas.

4.3 Realização da tarefa

Para se investigar a leitura de revisores de textos profissionais, foi realizado um experimento de detecção de erros em textos jornalísticos, no qual os participantes deveriam ler e revisar os textos apresentados em uma tela de computador, clicando com o mouse nos erros/inadequações encontrados. Foi utilizado o rastreador ocular produzido pela SR Research, modelo EyeLink 1000, disponível no Laboratório de Psicolinguística da UFMG. Esse rastreador é do tipo desktop, ou seja, a câmara e o iluminador ficam instalados abaixo do monitor de computador em que são projetados os estímulos, a uma distância de 40 a 70 cm da cabeça do sujeito. Foi utilizado um estabilizador de cabeça para minimizar os movimentos dos sujeitos e possibilitar o registro do reflexo pupilar e corneano. Embora o modelo de rastreador utilizado permita o registro binocular, foi realizado o registro de um dos olhos do sujeito, tendo sido definido qual deles antes do início da coleta de dados, em função da melhor calibração do equipamento.

O experimento foi realizado no Laboratório de Psicolinguística da Faculdade de Letras da UFMG. Os textos eram apresentados em uma tela de computador, aleatorizados. Primeiramente, eram apresentadas as instruções, o equipamento era calibrado e os participantes realizavam um pequeno treino de leitura e detecção de erros em três pequenos textos. Foi solicitado que os participantes lessem o texto procurando identificar

erros e inadequações e, ao identificarem um erro/inadequação, deveriam clicar com o mouse sobre ele. Além do registro do movimento ocular, foi obtida, também, uma medida cronométrica: o tempo de resposta. Além disso, constou nas instruções que os textos foram construídos especificamente para a pesquisa e, assim, os fatos narrados não precisavam necessariamente ser verdadeiros.

Caso algum participante pedisse informações mais detalhadas sobre a atividade de revisão, o experimentador se restringia a informar que era para ele se comportar como se estivesse fazendo a revisão do texto, com a detecção de erros e inadequações. Se perguntassem quais tipos de erros deveriam ser detectados, o experimentador informava que não poderia responder, destacando que eles deveriam marcar todos os erros que encontrassem. Tal procedimento foi adotado com vistas a verificar a definição que os sujeitos tinham da tarefa de revisão, por meio dos tipos de erros detectados.

Essa revisão de cada texto deveria ser realizada em, no máximo, 1 minuto/60segundos, pois, caso se ultrapassasse esse limite de tempo, o programa passava automaticamente para o próximo texto. Entre um texto e outro, havia ainda uma tela de *drift correction*, em que um círculo semelhante ao da calibração era projetado na tela no exato local em que se iniciaria o próximo texto. Como o sujeito deveria fixar esse círculo para que o experimentador passasse para o próximo texto, era garantido que, no início da leitura, o sujeito estaria com o olhar fixado na primeira palavra de cada texto. Isso evita variação entre textos e sujeitos quanto ao início da leitura de cada texto. A cor de fundo era branca, a letra preta em fonte Times New Roman de tamanho 20, com espaçamento 3,5. Cuidou-se de retirar o máximo possível o brilho da tela do computador, para evitar fadiga visual.

4.4 Análise dos dados

Foram analisadas diversas variáveis, de vários níveis: texto, sentença e trecho alvo. Assim, a análise dos dados se pautou em dois ângulos: o nível ao qual se aplica a variável investigada (texto, sentença ou trecho) e a relação entre movimento ocular e detecção de erros pelos sujeitos.

O nível de investigação da variável tem uma relação direta com o tipo de erro investigado: a supressão de preposição, por ser um erro que tem escopo local, requer medidas que se apliquem ao nível da palavra ou

do sintagma; já a anáfora incorreta tem escopo na integração de partes do texto e, assim, requer também medidas que correspondam aos níveis da sentença e do texto.

Além disso, neste trabalho há tanto as medidas do movimento ocular quanto a medida cronométrica de tempo de resposta. Por isso, foi possível investigar o movimento ocular e também a relação entre movimento ocular e detecção. Portanto, optou-se por utilizar medidas tradicionais nos estudos sobre leitura e movimento ocular, bem como por adaptar algumas medidas a esta pesquisa, conjugando-as com o tempo de resposta.

QUADRO 1 - Variáveis dependentes⁴

	Texto	Sentenças	Trecho
Movimento ocular	<ul style="list-style-type: none"> – Tempo total de leitura – Número de fixações nos textos – Duração média das fixações nos textos 	<ul style="list-style-type: none"> – <i>First-pass fixation time</i> nas sentenças alvo – Tempo total de fixação nas sentenças alvo – <i>Second-pass fixation time</i> nas sentenças alvo – <i>Regression-path reading time</i> nas sentenças alvo 	<ul style="list-style-type: none"> – <i>First-pass fixation time</i> – Tempo total de fixação no trecho alvo – Número de fixações no trecho alvo – <i>Regression-path</i>
Movimento ocular e detecção	<ul style="list-style-type: none"> – Tempo total de leitura até a detecção – Tempo total de leitura após a detecção – Número de fixações até a detecção – Número de fixações após a detecção – Duração média das fixações até a detecção – Duração média das fixações após a detecção 	<ul style="list-style-type: none"> – Tempo total de fixação nas sentenças alvo até a detecção – Tempo total de fixação nas sentenças alvo após a detecção – <i>Second-pass fixation time</i> nas sentenças alvo até a detecção – <i>Second-pass fixation time</i> nas sentenças alvo após a detecção 	<ul style="list-style-type: none"> – Tempo total de fixação no trecho alvo até a detecção – Tempo total de fixação no trecho alvo após a detecção

⁴ Neste trabalho, optou-se por manter a terminologia em inglês, uma vez que não há tradução já consagrada em português.

Nas medidas agrupadas como “Movimento ocular”, considerou-se tanto os textos em que os erros foram detectados quanto aqueles em que eles não foram detectados, de maneira a se comparar a detecção e a não detecção. Já nas medidas agrupadas como “Movimento ocular e detecção”, foram considerados somente os textos em que o erro foi detectado, uma vez que o momento da detecção (no qual o participante clicava com o mouse no erro) foi utilizado como um divisor para se analisar cada medida até e após essa detecção: por exemplo, o tempo total de leitura até a detecção do erro e o tempo total de leitura após a detecção do erro, medidas que correspondem ao tempo total de leitura até o momento em que o participante clica com o mouse sobre o erro e ao tempo total de leitura posterior a esse momento, até ele finalizar a tarefa.

O tempo total de leitura corresponde ao tempo total que o sujeito demorou lendo e revisando cada texto. Já o tempo total de leitura até a detecção do erro corresponde ao tempo que o sujeito demorou na leitura do texto até clicar com o mouse sobre o erro; e o tempo total de leitura após a detecção,⁵ ao tempo que o sujeito demorou desde que clicou com o mouse sobre o erro até terminar a revisão do texto. A mesma divisão aplica-se ao número de fixações no texto até e após a detecção do erro.

A duração média das fixações equivale à razão entre o tempo total de leitura, incluindo fixações e sacadas, e o número de fixações no texto. Assim, a duração média das fixações até a detecção do erro corresponde à razão entre o tempo total de leitura até a detecção e o número de fixações até a detecção. A mesma lógica aplica-se à duração média das fixações após a detecção do erro.

O *first-pass fixation time* corresponde à soma das durações de todas as fixações realizadas (seja na sentença alvo, quando no nível da sentença, ou na palavra alvo, quando no nível local) até seu abandono para a esquerda ou para a direita e durante a primeira leitura, ou seja, é uma medida que indica o tempo demandado em uma primeira leitura do alvo.

Já o tempo total de fixação nas sentenças equivale à soma de todas as fixações realizadas na sentença. Esse tempo também foi dividido em

⁵ Embora já fosse esperada diferença significativa entre as condições experimentais quanto às medidas de tempo após a detecção do erro, em função da localização de cada trecho alvo (na primeira sentença no caso da supressão de preposição e na terceira para a anáfora incorreta), tais medidas foram investigadas para se analisar se os grupos fariam releituras do texto e o engajamento com a tarefa de revisão.

duas partes: até e após a detecção do erro, seguindo a mesma lógica de contabilização apresentada acima.

O *second-pass fixation time* diz respeito à diferença entre o tempo total de fixação na sentença e o *first-pass fixation time*, ou seja, equivale ao tempo de retorno à sentença depois de ela ter sido lida uma primeira vez. Para essa medida, também foi contabilizado o valor até e após da detecção do erro.

A última medida do nível da sentença, o *regression-path reading time* corresponde à soma de todas as fixações realizadas pelo sujeito desde o momento em que ele fixa pela primeira vez a sentença alvo até quando ele faz uma fixação em alguma parte posterior do texto, contabilizando, além das fixações na própria sentença alvo, também as possíveis regressões realizadas a partes anteriores do texto.

O tempo total de fixação no trecho alvo diz respeito à soma de todas as fixações feitas nele, sendo que foi verificada, também, a soma de todas as fixações até a detecção do erro e após a detecção.

O *regression-path* corresponde à soma de todas as fixações realizadas pelo sujeito desde o momento em que ele fixa pela primeira vez o trecho alvo até quando ele faz uma fixação em alguma parte posterior do texto, contabilizando, além das fixações no próprio trecho alvo, também as possíveis regressões realizadas a partes anteriores do texto.

Foram variáveis independentes na análise do movimento ocular nos textos experimentais:

1. Tipos de erros:
 - a. supressão de preposição
 - b. anáfora incorreta
2. Grupos de revisão de textos:
 - a. revisores
 - b. não revisores
3. Detecção:
 - a. erros detectados
 - b. erros não detectados

O cruzamento dessas variáveis gera 8 grupos:

1. revisores – supressão de preposição – erro detectado;
2. não revisores – supressão de preposição – erro detectado;
3. revisores – supressão de preposição – erro não detectado;
4. não revisores – supressão de preposição – erro não detectado;
5. revisores – anáfora incorreta – erro detectado;
6. não revisores – anáfora incorreta – erro detectado;
7. revisores – anáfora incorreta – erro não detectado;
8. não revisores – anáfora incorreta – erro não detectado;

Primeiramente, os dados foram tabulados e explorados descritivamente, verificando-se medidas de tendência central como média, mediana e desvio padrão, bem como com a visualização por meio de gráficos como boxplot, gráfico de barras, gráfico de densidade e de médias, escolhidos conforme o tipo de variável investigada.

Foi, então, realizada a estatística inferencial, escolhendo-se o teste estatístico mais adequado em função dos dados analisados. O nível de significância adotado foi de 95%, portanto, α foi definido como 0,05. O programa utilizado para a análise estatística foi o R.⁶

Nas variáveis em que as amostras seguiam uma distribuição normal, foi escolhido o teste paramétrico. Como o experimento foi realizado com medidas repetidas, o teste paramétrico definido foi o modelo de regressão linear misto.⁷

Já nas variáveis em que as amostras não seguiam uma distribuição normal, os testes não paramétricos escolhidos foram o de Kruskal-Wallis ou de Wilcoxon-Mann-Whitney, tendo em vista tratar-se de testes de postos e, assim, não ser necessário que se conheça a distribuição das amostras.

5. Resultados

Os revisores detectaram 191 dos 279 erros, o que corresponde a 68,5% dos erros. Já os não revisores detectaram 165 dos 280 erros,

⁶ Disponível em <<https://www.r-project.org/>>.

⁷ Este teste foi escolhido, também, porque permite analisar a relação entre as variáveis ao longo da execução da tarefa, indicando aspectos importantes do comportamento dos participantes, não explorados neste artigo, mas disponíveis em Leite (2014).

o que equivale a 58,9%. Portanto, os revisores foram mais proficientes e detectaram mais erros do que os não revisores, mas, considerando somente os textos experimentais, essa diferença não foi estatisticamente significativa.

Nas tabelas apresentadas, serão indicados os valores das médias e o t ou p valor (conforme o teste estatístico). Informações mais detalhadas, como número de valores considerados por variáveis, desvio-padrão, graus de liberdade e gráficos podem ser consultadas em Leite (2014). Os códigos utilizados em todas as tabelas apresentadas serão os seguintes:

rev: revisores

nrev: não revisores

det: erro detectado

ndet: erro não detectado

sp: supressão de preposição

an: anáfora incorreta

N/A: não se aplica

*: diferença estatisticamente significativa

. : diferença marginalmente significativa

5.1 Nível do texto

A seguir, são apresentadas as variáveis em que foram detectados resultados estatisticamente significativos.

Nas variáveis em que foram consideradas as medidas até e após a detecção do erro, não foram apresentados os resultados estatisticamente significativos encontrados entre os tipos de erros, ou seja, quando se comparou a anáfora incorreta e a supressão de preposição, uma vez que essas diferenças já eram esperadas em função da composição dos textos (no primeiro caso, o erro encontra-se na terceira sentença e, no segundo, na primeira). Porém, nesses casos, foram comparados os grupos de revisão em um mesmo tipo de erro (revisores na anáfora x não revisores na anáfora e revisores na supressão de preposição x não revisores na supressão de preposição).

TABELA 1 – Resumo dos valores médios dos resultados obtidos nas variáveis do nível do texto – modelos de regressão linear mistos

Medida	Média		t valor	Média		t valor	Média		t valor
	rev	nrev		an	sp		det	ndet	
tempo total de leitura do texto	35473ms	27402ms	*rev>nrev 0,002	32709ms	30147ms	*an>sp 0,04	31845ms	30697ms	*det>ndet 0,01
número de fixações no texto	134	111	rev>nrev 0,06	129	117	*an>sp 0,002	123	123	-
duração média das fixações no texto	225ms	205ms	rev>nrev 0,06	212ms	218ms	*sp>an 0,000	218ms	210ms	det>ndet 0,05
tempo total de leitura após a detecção do erro	21590ms	13812ms	*rev>nrev 0,001	13323ms	20696ms	*sp>an 0,000	N/A	N/A	N/A
duração média das fixações até a detecção do erro	245ms	222ms	rev>nrev 0,06	219ms	243ms	*sp>an 0,000	N/A	N/A	N/A

TABELA 2 – Resumo dos resultados de interação obtidos nas variáveis do nível do texto – modelos de regressão linear mistos

Medida	Variável 1	Variável 2	Média (ms)	Interações (p valor)
Número de fixações no texto	det	sp	104	* an det > sp det (0,003) * an det > an ndet (0,01)
		an	130	
	ndet	sp	113	
		an	112	
Duração média das fixações no texto	det	sp	214	* sp det > an det (0,000) * sp det > an ndet (0,03) * sp det > sp ndet (0,000)
		an	210	
	ndet	sp	198	
		an	207	

TABELA 3 – Resumo dos resultados obtidos no nível do texto – comparações múltiplas do teste de Wilcoxon-Mann-Whitney com correção de Bonferroni

Medida	Variável 1	Variável 2	Mediana	Interações (p valor)
tempo total de leitura até a detecção do erro	sp	rev	7469 ms	* rev an > nrev an (0,000)
		nrev	7645 ms	
	an	rev	22301ms	
		nrev	17094 ms	
número de fixações até a detecção do erro	sp	rev	29	* rev an > nrev an (0,005)
		nrev	34	* nrev sp > rev sp (0,01)
	an	rev	104	
		nrev	81	
duração média das fixações após a detecção do erro	sp	rev	276 ms	* rev sp > nrev sp (0,02)
		nrev	265 ms	
	an	rev	377 ms	
		nrev	377 ms	
número de fixações após a detecção do erro	sp	rev	81,5	* rev an > nrev an (0,002)
		nrev	56,5	* rev sp > nrev sp (0,000)
	an	rev	38	
		nrev	20	

Na comparação entre os dois tipos de erros, verificou-se que os sujeitos tiveram um tempo maior de leitura quando houve a anáfora incorreta, o que está associado ao número de fixações nos textos, já que, também no que se refere a essa variável, houve diferença estatisticamente significativa entre as duas condições, com maior número de fixações nos textos em que houve a anáfora incorreta. Ao contrário, as fixações foram em média mais longas nos textos em que havia a supressão de preposição, o que pôde ser evidenciado também na duração média das fixações até a detecção do erro.

Além disso, na interação entre essas variáveis (tabela 3), os resultados indicam uma diferença entre a detecção da anáfora incorreta e a detecção da supressão de preposição: enquanto na anáfora incorreta o tempo total de leitura maior é relacionado ao número maior de fixações no texto, na supressão de preposição a detecção está relacionada a fixações em média mais longas. Essa diferença de padrão, portanto, pode estar relacionada a dois tipos diferentes de leitura: uma leitura mais voltada à superfície, com fixações mais longas, e uma leitura voltada a níveis mais globais, com mais fixações no texto.

No que se refere à análise depreendida no nível do texto, os resultados indicam que os revisores, em geral, fazem uma leitura mais demorada do texto, associada a um número maior de fixações.

Na comparação entre revisores e não revisores, os primeiros tiveram valores mais elevados no tempo total de leitura e tempo total de leitura após a detecção do erro (tabela 1), o que indica que, além de demandarem mais tempo lendo os textos, eles ainda mantiveram a leitura por mais tempo mesmo após terem detectado o erro. Quanto ao tempo total de leitura até a detecção do erro, quando comparados os grupos de revisão em um mesmo tipo de erro, os revisores apresentaram valores mais elevados do que os não revisores quando houve anáfora incorreta, mas não diferiram com relação à detecção da supressão de preposição (tabela 3).

Embora tenha havido uma diferença somente marginalmente significativa entre revisores e não revisores no que se refere ao número de fixações nos textos (tabela 1), ao se analisar esse número até a detecção do erro, na supressão de preposição, ele foi maior no grupo dos não revisores e, ao contrário, na anáfora incorreta, foi em média maior no grupo dos revisores. Após a detecção do erro, revisores fizeram mais fixações, independente do tipo de erro.

Quanto à duração média das fixações no texto, houve uma diferença marginalmente significativa entre revisores e não revisores, assim como quando foi analisada essa variável até a detecção do erro (tabela 1), sendo que a tendência foi de os revisores fazerem fixações em média mais longas. Houve diferença estatisticamente significativa entre os grupos no que se refere à duração média das fixações após a detecção do erro (tabela 3).

Dessa forma, os revisores mostraram um padrão de leitura diferente para cada tipo de erro: na detecção da supressão de preposição, as fixações tenderam a ser mais longas (duração média das fixações marginalmente significativa), e na da anáfora incorreta, o número de fixações foi maior.

O tempo total de leitura também foi maior nos textos em que o erro foi detectado do que naqueles em que o sujeito não detectou o erro, assim como o número de fixações, porém, tal achado pode decorrer da própria tarefa, uma vez que, nos textos em que o erro foi detectado, o participante demandou tempo clicando com o mouse no erro.

5.2 Nível da sentença

Quanto às medidas apresentadas no nível da sentença, na supressão de preposição, considerou-se somente a primeira sentença, na qual estava localizado o erro. Já na anáfora incorreta, foram analisadas tanto a primeira quanto a terceira sentença, uma vez que nelas estavam localizados, respectivamente, o referente e a anáfora incorreta.

TABELA 4 – Resumo dos resultados obtidos nas variáveis do nível da sentença – modelos de regressão linear mistos

Medida	Média (ms)		t valor	Média (ms)		t valor	Média (ms)		t valor
	rev	nrev		an	sp		det	ndet	
tempo total de fixação na primeira sentença	11274	9227	*rev>nrev 0,02	9971ms	10527	0,45	10592	9646	*det>ndet 0,000
tempo total de fixação na terceira sentença	11050	8240	*rev>nrev 0,009	N/A	N/A	N/A	10756	8655	*det>ndet 0,000
tempo total de fixação na primeira sentença até a detecção do erro	9075	8270	*rev>nrev 0,008	9378	8305	0,24	N/A	N/A	N/A
tempo total de fixação na terceira sentença até a detecção do erro	6341	4868	0,21	N/A	N/A	N/A	N/A	N/A	N/A

TABELA 5 – Resumo dos resultados de interação obtidos nas variáveis do nível da sentença – modelos de regressão linear mistos

Medida	Variável 1	Variável 2	Média (ms)	Interações (p valor)
tempo total de fixação na primeira sentença até a detecção do erro	sp	rev	8192	* rev an > nrev an (0,03) * rev an > rev sp (0,03)
		nrev	8431	
	an	rev	10522	
		nrev	7981	

TABELA 6 – Resumo dos resultados obtidos no nível da sentença
– comparações múltiplas do teste de Wilcoxon-Mann-Withney
com correção de Bonferroni

Medida	Variável 1	Variável 2	Variável 3	Mediana (ms)	Interações (p valor)
<i>second-pass fixation time</i> na primeira sentença	det	sp	rev	2241	* rev an det > nrev an det (0,000)
			nrev	182,5	* rev an det > nrev sp det (0,000)
		an	rev	5076	* rev an det > rev sp det (0,000)
			nrev	2088	* rev an det > nrev an ndet (0,000)
	ndet	sp	rev	4274	* rev an det > nrev sp ndet (0,000)
			nrev	1846	* rev sp det < nrev sp det (0,000)
		an	rev	1668,5	* nrev sp det < nrev an ndet (0,03)
			nrev	2092	* nrev sp det < rev an ndet (0,03)
<i>first-pass fixation time</i> na terceira sentença	an	det	rev	5338	* nrev det > nrev ndet (0,000)
			nrev	6181	* nrev det > rev ndet (0,000)
		ndet	rev	6072	* rev det < nrev ndet (0,006)
			nrev	5739	
<i>second-pass fixation time</i> na terceira sentença	an	det	rev	8659	* rev det > nrev det (0,000)
			nrev	6039	* rev det > nrev ndet (0,000)
		ndet	rev	2118,5	* rev det > rev ndet (0,000)
			nrev	1666	* nrev det > nrev ndet (0,000)
<i>regression-path reading time</i> na terceira sentença	an	det	rev	4223,5	* rev det > nrev ndet (0,000)
			nrev	4282	* rev det > rev ndet (0,000)
		ndet	rev	2489,5	* nrev det > rev ndet (0,002)
			nrev	2199	* nrev det > nrev ndet (0,000)
tempo total de fixação na primeira sentença após a detecção do erro	det	sp	rev	2208,5	* rev an > nrev an (0,001)
			nrev	525,5	* rev an < rev sp (0,000)
		an	rev	506,0	* nrev an < nrev sp (0,000)
			nrev	0	* nrev an < rev sp (0,000)
tempo total de fixação na terceira sentença após a detecção	det	an	rev	5232	* rev > nrev (0,000)
			nrev	2933	
<i>second-pass fixation time</i> na primeira sentença até a detecção	det	sp	rev	0	* rev an > nrev an (0,02)
			nrev	0	* rev an > rev sp (0,000)
		an	rev	2904,5	* rev an > nrev sp (0,000)
			nrev	1526	* nrev an > rev sp (0,000)

Quando foram comparados revisores e não revisores, houve diferença estatisticamente significativa quanto ao tempo total de fixação na primeira e na terceira sentenças do texto, quanto ao tempo total de fixação na primeira sentença do texto até a detecção do erro, bem como quanto ao tempo total de fixação na primeira e na terceira sentenças após

a detecção do erro. Em todas essas situações, seja na anáfora incorreta ou na supressão de preposição, os revisores apresentaram valores em média mais elevados do que os não revisores. Os dois grupos também diferiram quanto ao *second-pass fixation time* na primeira e na terceira sentença, sendo que os revisores tiveram um tempo de retorno a essas sentenças maior do que os não revisores. Os revisores também tiveram um tempo de retorno à primeira sentença até a detecção da anáfora incorreta (*second-pass fixation time* até a detecção do erro) maior do que os não revisores, mas os dois grupos não diferiram no que se refere à supressão de preposição, pois, em ambos, a detecção da supressão de preposição ocorreu na primeira leitura da sentença.

5.3 Nível local

Na condição de supressão de preposição, foi analisado somente o tempo total de fixação no trecho alvo, que abarca a 15^a e a 16^a palavra do texto. Já na condição de anáfora incorreta, foi analisado o tempo total de fixação no referente e na anáfora, em separado.

TABELA 7 – Resumo dos resultados obtidos nas variáveis do nível local – modelos de regressão linear mistos

Medida	Média		t valor	Média		t valor	Média		t valor
	rev	nrev		an	sp		det	ndet	
número de fixações no alvo – sp	10,74	11,74	0,16	N/A	N/A	N/A	11,91	7,68	*det>ndet 0,000
tempo total de fixação no referente até a detecção	1169ms	831ms	*rev>nrev 0,007	N/A	N/A	N/A	N/A	N/A	N/A

TABELA 8 – Resumo dos resultados de interação obtidos nas variáveis do nível local – modelos de regressão linear mistos

Medida	Variável 1	Variável 2	Média	Interações (p valor)
número de fixações no alvo – sp	det	rev	11	* rev det > rev ndet (0,000) * rev det > nrev ndet (0,000) * nrev det > rev ndet (0,000) * nrev det > rev ndet (0,000)
		nrev	13	
	ndet	rev	8	
		nrev	7	

TABELA 9 – Resumo dos resultados obtidos no nível local
– comparações múltiplas do teste de Wilcoxon-Mann-Withney
com correção de Bonferroni

Medida	Variável 1	Variável 2	Mediana	Interações (p valor)
tempo total de fixação no alvo – sp	det	rev	2980ms	* nrev det > nrev ndet (0,000)
		nrev	3233ms	* nrev det > rev ndet (0,000)
	ndet	rev	1072ms	* rev det > nrev ndet (0,000)
		nrev	1085ms	* rev det > rev ndet (0,000)
tempo total de fixação no referente	det	rev	1229ms	* rev det > nrev det (0,002)
		nrev	726ms	* nrev det > nrev ndet (0,008)
	ndet	rev	724,5ms	* rev det > nrev ndet (0,000)
		nrev	497ms	* rev det > rev ndet (0,000)
tempo total de fixação na anáfora	det	rev	2092ms	* rev det > nrev ndet (0,000)
		nrev	1799ms	* rev det > rev ndet (0,000)
	ndet	rev	669ms	* nrev det > nrev ndet (0,000)
		nrev	475ms	* nrev det > rev ndet (0,000) * rev det > nrev det (0,06)
número de fixações no referente	det	rev	6	* nrev det > nrev ndet (0,003)
		nrev	4	* rev det > nrev ndet (0,000)
	ndet	rev	3	* rev det > rev ndet (0,000)
		nrev	3	* rev det > nrev det (0,05)
número de fixações na anáfora	det	rev	8	* nrev det > nrev ndet (0,000)
		nrev	7	* rev det > nrev ndet (0,000)
	ndet	rev	3	* rev det > rev ndet (0,000)
		nrev	2	* nrev det > rev ndet (0,000) * rev det > nrev det (0,05)
tempo total de fixação no alvo após a detecção – sp	det	rev	675,5ms	* rev > nrev (0,000)
		nrev	104ms	
tempo total de fixação na anáfora após a detecção	det	rev	246,5ms	* rev > nrev (0,000)
		nrev	0ms	

Quando foi verificado o tempo total de fixação no trecho alvo durante toda a leitura do texto, sem se levar em consideração o tempo da detecção, não foram encontradas diferenças estatisticamente significativas entre os revisores e não revisores. Ao contrário, houve diferenças estatisticamente significativas entre os textos em que os erros foram detectados e aqueles em que os erros não foram detectados,

sendo que, quando houve a detecção, o tempo total de fixação no alvo foi maior. Esse padrão foi encontrado nos textos em que havia supressão de preposição e também naqueles em que havia a anáfora incorreta. Quanto a esse tipo de erro, somente no referente houve diferença entre revisores e não revisores, com valores mais elevados entre os primeiros. Os mesmos resultados foram obtidos na análise do número de fixações no alvo, mas, nesse caso, a diferença entre revisores e não revisores quanto ao número de fixações no referente e na anáfora foi apenas marginalmente significativa.

Da mesma forma, ao se analisar o tempo total de fixação no alvo até a detecção do erro, ou seja, a soma de todas as fixações realizadas no alvo até que o participante clicasse no mouse, inclusive os retornos feitos na palavra até então, houve diferença significativa entre revisores e não revisores somente quanto ao referente, na anáfora incorreta, sendo que, nesse caso, os valores em geral mais elevados foram encontrados entre os revisores. Na análise do tempo total de fixação no referente após a detecção, ou seja, depois de terem clicado com o mouse no erro, porém, não houve diferença significativa entre revisores e não revisores, já que ambos os grupos, na maioria das vezes, não retornaram ao referente.

Já no que se refere à região da anáfora incorreta, houve diferença estatisticamente significativa entre revisores e não revisores quanto ao tempo total de fixação após a detecção do erro. Portanto, os revisores demandaram mais tempo fixando o referente para detectarem a incongruência e, possivelmente, buscar resolvê-la. Depois de detectarem o erro, eles ainda fizeram mais retornos à anáfora, antes de finalizarem a leitura do texto.

Na supressão de preposição, como ocorreu com o tempo total de fixação no alvo, também não houve diferença estatisticamente significativa entre revisores e não revisores quanto ao tempo total de fixação no alvo até a detecção do erro. O oposto ocorreu quanto ao tempo total de fixação no alvo após a detecção do erro: os revisores tiveram valores, em média, mais elevados do que os não revisores.

Dessa forma, embora não tenha havido diferença significativa entre os revisores e os não revisores ao se analisar o tempo total de fixação no alvo na leitura do texto, essa diferença surgiu ao se analisar o movimento ocular levando-se em conta o tempo de detecção do erro. Nesse caso, os revisores, em geral, tiveram valores mais elevados do que os não revisores e, nos textos em que os erros foram detectados, os

valores de tempo total de fixação no alvo foram mais elevados do que nos textos em que os erros não foram detectados. Essa variável, portanto, é relevante ao se investigar a proficiência na detecção dos erros.

6. Considerações finais

Os resultados indicaram que, em geral, os textos em que houve a anáfora incorreta apresentaram valores mais elevados nas medidas investigadas, em especial do nível do texto, inclusive com um tempo total de leitura maior, relacionado com um maior número de fixações. Esse resultado corrobora os estudos da área, indicando que erros de níveis que exigem a integração de partes dos textos acarretam mais dificuldade no processamento. A exceção foi a duração média das fixações no texto e, também, até a detecção do erro. Nesse caso, os valores foram mais elevados nos textos em que havia a supressão de preposição. Como nessa medida a interação entre o tipo de erro e a detecção foi significativa, pode-se dizer que o aumento da duração média das fixações no texto, inclusive no início, até a detecção, favorece a identificação da supressão de preposição. Tendo em vista a localização dos erros nos textos e a complexidade deles, essas diferenças podem decorrer das estratégias globais utilizadas na leitura numa fase inicial ou mais tardia. Por isso, é importante que outros estudos explorem esse achado inicial, inclusive comparando uma condição de base, com e sem erro.

Os resultados desta pesquisa corroboram o proposto por Vauras, Hyona e Niemi (1992), segundo os quais as incoerências aumentam tanto o tempo de fixação quanto o número de regressões. Esta pesquisa, portanto, confirmou dados da área de Linguística que diferenciam uma leitura voltada para níveis mais globais de uma leitura voltada para níveis mais superficiais, contribuindo, também, para a caracterização desses tipos de leitura, assim como para mostrar quais estratégias podem ser mais indicadas para cada nível. Nesse sentido, um achado importante, porém inicial, foi que a detecção da supressão de preposição é favorecida quando são realizadas fixações em média mais longas. Muitos estudos da área propõem que a leitura voltada para níveis superficiais deve ser diferente daquela voltada para níveis mais globais, sem indicarem, porém, como essa diferença se processa quanto ao movimento dos olhos. A duração mais longa das fixações parece ser um parâmetro significativo, sendo necessários outros estudos, inclusive com diferentes tipos de erros de superfície, para confirmar esse achado.

Com relação às diferenças entre os textos em que o erro foi detectado e aqueles em que o erro não foi detectado, a perspectiva geral é de que a detecção do erro é acompanhada pelo aumento do tempo de fixação, seja no nível do texto, da sentença ou local. No nível local, nos textos em que o erro é detectado, o tempo total de fixação no alvo é maior, o que ocorre tanto na supressão de preposição quanto na anáfora incorreta. Tal diferença, porém, pode ser decorrente do tempo demandado pelos participantes para clicar com o mouse no texto.

Por fim, no que se refere à diferença entre revisores e não revisores, a perspectiva geral foi de que os revisores apresentaram valores mais elevados nas variáveis em que houve diferença significativa, o que significa que eles foram mais lentos na leitura do que os não revisores. Isso ocorreu no nível do texto, da sentença e no nível local.

Os resultados, portanto, confirmaram a hipótese inicial de que os revisores profissionais teriam uma leitura mais demorada, menos automática, mais detalhada e controlada. De acordo com Hayes (2004), com a prática extensiva, certos aspectos da edição podem se tornar automáticos, mas não a atividade de edição como um todo, mesmo em sujeitos com muita experiência, motivo pelo qual o controle da leitura é um pressuposto da revisão. Klein e Hoffman (1992), ao discorrerem sobre a *expertise*, destacam que a performance diferenciada de um *expert* pode ser identificada por diversos fatores: variabilidade-consistência, acurácia, completude e velocidade. Quanto à velocidade, os autores destacam que, em alguns casos, *experts* podem ser mais demorados do que novatos no cumprimento de uma tarefa, inclusive quando os novatos, por impulsividade, adotam um comportamento muito rápido. Esta pesquisa indica a ocorrência desse tipo de situação.

Os revisores profissionais também foram mais específicos quanto aos retornos realizados. Nos textos em que havia a anáfora incorreta, tiveram um tempo total de fixação na primeira sentença até a detecção do erro maior do que os não revisores e, também, um tempo total de fixação maior no referente. Como o comportamento geral dos sujeitos, ao se depararem com a anáfora incorreta, foi retornar a partes anteriores do texto, os revisores foram mais específicos e fixaram por mais tempo a parte do texto que era realmente importante para a resolução do problema, ou seja, o referente e a sentença na qual ele está contido. Além disso, após detectarem a anáfora incorreta, os revisores fizeram retornos mais longos à terceira sentença do texto e, também, à anáfora em si, o que

indica que, também neste caso, eles foram mais específicos. De acordo com Hyona e Nurminen (2006), leitores proficientes tendem a direcionar suas regressões a partes realmente informativas do texto.

Por fim, deve-se considerar que esta pesquisa investigou especificamente a detecção de erros, que é uma das etapas da revisão, sendo que, além dela, é necessário também realizar as correções, etapa esta que não foi investigada neste momento, mas que merece consideração em estudos posteriores da área, inclusive naqueles em que se verifique a qualidade das correções realizadas, de cunho, portanto, qualitativo. Tendo em vista que os estudos sobre a revisão de textos profissional são escassos, principalmente com abordagem cognitiva, essa pesquisa teve um caráter exploratório e trouxe considerações iniciais importantes não só para o estudo da performance de revisores profissionais, mas também para o estudo do processo de leitura e revisão de textos.

Referências

ALAMARGOT, D. *et al.* Eye and Pen: a new device to study reading during writing. *Behaviour Research Methods, Instruments and Computers*, Bethesda, USA, v. 38, n. 2, p. 287-299, 2006. doi.org/10.3758/BF03192780.

HAYES J. R. A new framework for understanding cognition and affect in writing. In: C. Levy; S. Ransdell (Ed.). *The science of writing: theories, methods, individual differences and applications*. Mahwah: Lawrence Erlbaum Associates, 1996. p. 1-28.

HAYES, J. R. *What triggers revision?* In: ALLAL, L.; CHANQUOY, L.; LARGY, P. (Ed.). *Revision: cognitive and instructional processes*. Kluwer Academic Publishers, Massachusetts, 2004, p. 9-22. doi.org/10.1007/978-94-007-1048-1_2.

HAYES J. R. *et al.* Cognitive processes in revision. In: ROSENBERG, S. (Ed.). *Advances in psycholinguistics*: Cambridge: Cambridge University Press, 1987. v. II: Reading, writing, and language processing, p. 176-240.

HAYES, J. R.; FLOWER, L. S. Identifying the organization of writing processes. In: GREGG, L. W.; STEINBERG, E. R. (Ed.). *Cognitive processes in writing: An interdisciplinary approach*. Lawrence Erlbaum Associates, Hillsdale, 1980. p. 3-30.

HEURLEY, L. La révisión de texto: L'approche de la Psychologie Cognitive. *Langages*, Larousse, v. 164, n. 4, p. 10-25, 2006. doi.org/10.3917/lang.164.0010.

HYONA, Jukka; KAAKINEN, Johanna K.; LORCH, Robert F. Jr. Individual Differences in Reading to Summarize Expository Text: Evidence From Eye Fixation Patterns. *Journal of Educational Psychology*, Taylor & Francis, v. 94, n. 1, p. 44-55, 2002. doi.org/10.1037/0022-0663.94.1.44.

HYONA, Jukka; NIEMI, Pekka. Eye Movements During Repeated Reading of a Text. *Acta Psychologica*, Elsevier, v. 73, p. 259-280, 1990. doi.org/10.1016/0001-6918(90)90026-C.

HYONA, Jukka; NURMINEN, Anna-Mari. Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology*, John Wiley & Sons, v. 97, p. 31-50, 2006. doi.org/10.1348/000712605X53678.

KLEIN, G. A.; HOFFMAN, R. R. Seeing the invisible: perceptual-cognitive aspects of expertise. In: RABINOWITZ, M. (Ed.). *Cognitive Science Foundations of Instruction*. Mahwah, N. J: Erlbaum, 1992. p. 203-226.

LAND, Michael F. Fixation Strategies during active behavior: a brief history. In: VAN GOMPEL, R. P. G. *et al.* (Ed.). *Eye Movements: a window on mind and brain*. Amsterdã: Elsevier, 2007. p. 76-95.

LEITE, Délia Ribeiro. *O olhar do profissional: estudo do movimento ocular na leitura realizada por revisores de textos*. 2014. Tese (Doutorado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

LUEGI, Paula. *O registro do movimento dos olhos durante a leitura de textos*. Dissertação de mestrado. Universidade de Lisboa, 2006.

RAYNER, Keith. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, Washington, v. 124, n. 3, p. 372-422, 1998. doi.org/10.1037/0033-2909.124.3.372.

STAUB, Adrian; RAYNER, Keith. Eye movements and on-line comprehension processes. In: GASKELL, G. (Ed.). *The Oxford Handbook of Psycholinguistics*. Oxford, UK: Oxford University Press, 2007. p. 327-342.

VAURAS, Marja; HYONA, Jukka; NIEMI, Pekka. Comprehending coherent and incoherent texts: evidence from eye movement patterns and recall performance. *Journal of Research in Reading*, Wiley online Library, v. 15, n. 1, p. 39-54, 1992. doi.org/10.1111/j.1467-9817.1992.tb00020.x.

WAGNER, Luiz Roberto; CUNHA, Djenane S. W. *Categorias de um revisor*. *Língua Portuguesa – conhecimento prático*, 2012. n. 37, p. 12-13.

Revisers' eye movement in reading

O movimento ocular na leitura realizada por revisores de textos

Délia Ribeiro Leite

Conselho Regional de Contabilidade de Minas Gerais, Belo Horizonte, Minas Gerais / Brasil
dribeiroleite@yahoo.com.br

José Olímpio de Magalhães

Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais / Brasil
joseolimpiomagalhaes@yahoo.com.br

Abstract: This study aims to investigate the reading processing of professional revisers, comparing their eye movement with subjects who do not work professionally in revision (non-revisers). In the reading and error detection task proposed and performed by 14 revisers and 14 non-revisers, the participants had to read and edit short journalistic texts displayed on a computer, by clicking on the errors and/or inadequacies they found using the mouse. There were two types of errors: a) missing preposition and b) incoherence generated by an incorrect nominal anaphora. In the statistical analysis, when normality was ensured, mixed models were carried out; otherwise, the analysis was carried out by means of non-parametric tests. Considering the eye movement analysis, the general perspective indicates that revisers presented the highest values, meaning they were slower in reading than non-revisers, which occurred concerning the measures of the text, sentence, and local levels. This research contributes to a characterization of reading processes involved in the revision of texts by professional revisers.

Keywords: revision; professional revisers; eye movement; proficiency.

Resumo: Este estudo objetivou investigar o processamento da leitura de profissionais revisores de textos, comparando o movimento ocular destes com o de sujeitos que não trabalham profissionalmente com revisão (não revisores). Na tarefa de leitura e detecção de erros realizada por 14 revisores e 14 não revisores, os participantes deveriam ler e revisar pequenos textos jornalísticos projetados no computador, clicando com o mouse nos erros e/ou inadequações que encontrassem. Havia dois tipos de erros nos estímulos experimentais: a) supressão de preposição e b) incoerência gerada por uma anáfora nominal incorreta. Na análise estatística, quando garantida a normalidade, foram realizados modelos mistos; do contrário, a análise foi realizada por meio de testes não paramétricos. Quanto à análise do movimento ocular, a perspectiva geral foi de que os revisores apresentaram valores mais elevados, o que significa que eles foram mais lentos na leitura do que os não revisores, o que foi identificado nas medidas do nível do texto, da sentença e do trecho alvo. Esta pesquisa contribui para uma caracterização dos processos de leitura envolvidos em tarefas de revisão de textos por revisores profissionais.

Palavras-chave: revisão de textos; revisão profissional; movimento ocular; proficiência.

Recebido em: 9 de dezembro de 2016.

Aprovado em: 9 de janeiro de 2017.

1 Introduction

Revision is directly related to the text reading and production process. Therefore, when one rereads a text one has produced, one seeks to assume an exotopic view¹ and check different aspects of the textual structure

¹ The exotopic view is that in which reading is done by using a different gaze from that of the text's producer, a gaze that would stimulate the reading performed by the reader to which the text is intended.

² A *saccade* is a quick, simultaneous movement of both eyes between two phases of fixation in the same direction.

in order to identify possible errors, as well as segments whose interpretation poses difficulty to the reader. However, authors are not always the most suitable readers to revise their own texts, whether for technical reasons, related to the lack of specific knowledge in the language study field (in the case of authors in other areas), or for practical reasons: exotopy may be compromised, given that the text, when read by its own author, loses its unpredictability, which may cause language inadequacies to go unnoticed.

This is the reason for the existence of the professional text reviser, whose occupation is to read texts produced by others in order to identify language inadequacies and propose different corrections or changes.

Despite being an old and important activity, very few studies focus on professional revisers, although there is a perception that this task requires special qualification and demand, particularly regarding reading: “Revisers do not read as other men do, they photograph the word visually; and the texts are reflected in their corrections.” (WAGNER; CUNHA, 2012, p. 12).

Hence, this work aims at investigating how the reading performed by professional text revisers is processed. For such, the eye-tracking investigation method was used. Fixation and saccade² patterns performed by these professionals during revision were checked, as well as whether these patterns are similar to those detected when individuals who are not professional revisers read the same texts. Therefore, this work’s purposes are:

- To investigate eye movement patterns in reading performed by professional text revisers, in an error detection task;
- To compare the eye movement performed by professional text revisers to that of subjects who do not revise texts professionally, in a reading and error detection task;
- To determine if professional text revisers and subjects who do not revise texts professionally perform a reading directed both to the textual surface and to more global text levels.

This study contends that the reading performed by text revisers is less automatic, more controlled and more detailed, which leads to more proficiency in the revision activity.

2 Text revision in a cognitive approach

Heurley (2006) points out three major views of revision from the cognitive standpoint: revision as effective text change; revision as a subprocess of the writing process, aimed at improving the previously written text; and revision as a distinct component of the written production, which would involve a set of subprocesses and procedures implied in the control of writing.

Hayes et al. (1987), in turn, propose a model specifically for revision. The model is divided into processes and knowledge. The first process involved is defining the task, which includes specific issues, such as the reviser's intents, the text traits to be examined, and how the revision is to be conducted. The second process is the evaluation, when the reading objects are selected: understanding, evaluating, or detecting problems. Based on the representation made of the problem, a strategy will be selected, with the possibility for modifying or controlling the revision process in itself, or modifying the text. By modifying the revision process, the reviser may ignore the problem, seek more information to improve the diagnosis, or postpone the action, even when the reviser opts to do more than one reading, one of which directed to a higher level, and another one related to textual surface aspects. By contrast, if the reviser chooses to modify the text, he/she may rewrite it or revise it. In this case, the revision includes those cases in which the reviser corrects inadequacies, preserving the original text as much as possible. An important issue is that revisers may redefine the task as the revision is performed. Thus, the knowledge used, including goals, criteria, and restrictions, is dynamically modified during revision. Revision is related to reading for detecting problems, which, according to the authors, is different, for instance, from reading directed only to understanding.

In 1996, Hayes proposed Hayes and Flower's (1980) writing model be reformulated, with the main purpose of adding important cognitive elements, such as the working memory, as well as to reorganize existing elements to show the relationship between them and more general cognitive processes involved in several types of activities. Revision is therefore seen as a form of text interpretation. The author proposes a new revision model, in which there is a control structure, consisting of a revision task scheme; fundamental processes, including reflection, processing, and text producing; and the resources used, which are the working memory and

the long-term memory. The revision, along with the reflection and text production, is part of the subcomponent of cognitive processes, which, in turn, is part of a greater component. The latter, in addition to cognitive processes, involves an affective/motivational subcomponent, the working memory and the long-term memory. Thus, the author contends that, to understand the revision, one needs to understand not only the process, but also its control structure and the resources involved. The author points out that this set of requirements is acquired with practice, and this may also be inferred from the difference between new and experienced text producers and revisers, as explored by Hayes et al. (1987).

As shown, in these studies' scope, revision no longer plays a secondary role in the writing process, and revision is now seen as an important stage in this process, acting on several levels. Revision has been treated more and more as a control activity, which may operate separately or throughout the writing process.

3 Eye movement and reading

Eyes move in saccades alternating with fixations. While a saccade corresponds to a rapid eye movement to move the focus from one area to the other, a fixation is equivalent to the time spent focusing on a given area. It is possible to determine time and space variables when the eye movement is recorded, as there is a spatial displacement, that is, the point to where the movement is directed, and a temporal displacement, represented by the moment when this movement occurs. In addition, it is known that the type of movement depends heavily on the required information (LAND, 2007, p. 78).

Furthermore, the eye movement does not always occur in a linear manner, as there are situations in which the eye is drawn back to previously focused areas. According to Luegi (2006, p. 23), "Approximately 15% of saccadic movements during reading are regression movements, that is, they go from right to left, towards previous areas of the text, on the same line or a few lines above."

Regressions may be corrections of saccades that are too long and, therefore, the eyes must make a brief return. They may also result from the difficulty in understanding some part of the text. In this case, regressions are usually longer (longer than 10 characters on the same line or towards previous lines) and the behavior regarding the return varies according to the reader's proficiency.

Eye movement recording is used in a wide range of studies on language processing. Many variables may affect the values of fixations and saccades, including whether the reading is done silently or aloud. As there are no specific studies on eye movement during text revision, this section presents studies that address issues associated with the types of errors present in the experimental texts examined in this study, as well as factors taken into account in the analysis performed in this study, such as text rereadings and returns.

As the eye movement is processed in fixations and saccades, words may be skipped between fixations. It is well-known that functional words tend to be skipped, that is, not focused on. According to Rayner, functional words are fixated approximately 35% of the time (RAYNER, 1998, p. 375). This may result from them being more easily identifiable, that is, more predictable based on the context or even to the fact that they are more frequent (STAUB; RAYNER, 2007). The fact that they are short words may also affect their skipping, given that, as the word size increases, the probability of its fixation also increases – 2 to 3-letter words are fixated, in general, 25% of the times (RAYNER, 1998, p. 375).

Hyönä and Nurminen (2006) investigate regressions and, according to their results, readers who tend to direct the returns to parts of the text that are truly informative are able to provide a better summary of what they have read. Thus, there are also aspects related to higher textual levels that interfere in eye movements. Authors researched how competent readers read in order to confirm the classification proposed in a prior study (HYÖNÄ; KAAKINEN; LORCH, 2002) and to determine whether the readers themselves are aware of the reading strategies they use. Three reader patterns were identified: those who process topic structures, who were the minority, followed by the fast linear readers, and, finally, by the slow linear readers, who were the majority. The readers are aware of the time required for reading, whether they are slow or not, as well as whether they return or not to parts of the text; however, they do not have a precise idea of the places to which they comeback. Therefore, there are idiosyncratic differences related to eye movement and to reading, although the instructions provided in the task may affect the patterns. In the case of the research above, the readers were asked to summarize the text, which may have affected how they behaved.

Hyönä and Niemi (1990) investigate a text rereading. The rereading facilitates the process and, thus, all measurements investigated

are reduced, both the number and the time of fixations, as well as the number of returns. The authors also investigate whether the rereading affects the trend to fixate the segments more often with important information. In this sense, the conclusion reached is that the most important parts tend to be fixated more often, but this occurs both in the first reading and in the rereading. However, the rereading facilitates the processing of sentences that are more informative than of those that are not informative. Furthermore, for each successive reading, the average of fixations is longer at the beginning of texts than at the end.

Vauras, Hyönä, and Niemi (1992) investigate the reading of coherent and incoherent texts. In incoherent texts, the sentence order is changed. As their methodology, the authors use the recording of eye movement as well as an off-line measurement, which is an analysis of text rewriting performed by the readers. They conclude that incoherences increase the fixation time as well as the number of returns. Nevertheless, although in coherent texts the increased fixation time results in improved rewriting of the fixated part, the same is not true for incoherent texts.

Luegi (2006) researches the recording of eye movement among European Portuguese readers. The author presents modified sentences intended to create inconsistencies (ungrammaticality or ambiguity) and identifies the readers' reaction. In addition, these manipulations are done in two texts: one in which technical terms are predominant and another whose subject is more common. The author identified that the effect of variables is only visible when dealing with a text with technical vocabulary and sentence manipulation, thus generating inconsistencies. These results are seen both in the number of fixations and in the total reading time. Furthermore, the author determines that readers are quite precise regarding regressions and return specifically to the place where the problems are located.

Alamargot et al. (2006) record the eye movement and the graphomotor execution concomitantly, to investigate the writing process. Although the work focuses on text revision, they identified that the detection of typos can occur in very short times, as can those of a reading fixation. They can even occur in parallel with the graphomotor execution.

Therefore, the recording of eye movement may indicate important aspects of the reading processing. Thus, experimental works demonstrate that both the physiology of vision and the reading processing are determined regarding the eye movement pattern adopted by the subjects.

These studies will support this work's hypothesis that there will be a differentiated reading pattern according to the type of error, as well as to the subjects' experience with text revision.

4 Methodology

4.1 Participants

Subjects who participated in the research were divided into two groups: non-revisers and revisers, each group consisting of 14 subjects. The non-reviser group consisted of college students who did not major in Languages or Communications, as graduation in these two majors is a pre-requisite for investiture in several revision positions and, therefore, these students have a more in-depth qualification in language studies, which sets them apart from a group representing non-reviser subjects. To select the revisers, subjects were required to have worked professionally as revisers, in public agencies, in the private sector, or as freelancers. All revisers were Language majors, one of them was a senior in college, but had been working as a text reviser for three years. Regarding the experience time, 5 of them had worked as revisers for 1 to 4 years; 5 had worked between 5 and 10 years; and 4 had worked for more than 10 years.

4.2 Constructing stimuli

The texts, from the journalistic genre, written specifically for the experiment, consisted of an informative paragraph. The problems found in the experimental texts were of two types:

- missing preposition;
- incoherences generated by an incorrect nominal anaphora.

These types of problems were selected due to the processing cost and represent two different levels: the missing preposition is related to the word and phrase levels; while the incoherence generated by an incorrect nominal anaphora is related to the text level, as readers were expected to perform the integration between the text sentences.

In addition, texts without mistakes and with other types of errors, considered distracters, were inserted. In each experimental session, 20

experimental texts were presented, 10 of which contained a missing preposition, and 10 with incorrect anaphora, which did not include any other errors/inappropriacies whose detection was expected. In addition, 30 other distracter texts were included (8 of which without errors, 2 with one error, 7 with two errors, 9 with three errors, and 4 with four errors).²

In examples 1 and 2 below, two examples of experimental texts are presented, one of which with a missing preposition (presence of two teams' fan clubs) and the other with an incorrect anaphora (American referring to Argentinian).

Example 1 – Experimental text – missing preposition

Após um longo período, o clássico entre Atlético Mineiro e Cruzeiro contará com a presença duas torcidas mineiras. O Cruzeiro será o mandante da partida e vai disponibilizar uma parte dos ingressos para a torcida adversária. O jogo do Campeonato Mineiro será na próxima semana, na inauguração do novo Mineirão, e terá policiamento reforçado.
(After a long period, the classic game between Atlético Mineiro and Cruzeiro will receive the presence two teams' fans from Minas Gerais. Cruzeiro will be the home team and will provide part of the tickets to the opposing fans. The Minas Gerais championship game will take place next week, when the new Mineirão stadium will be opened, and police operation will be reinforced.)

Example 2 – Experimental text – incorrect anaphora

Na noite de ontem, um bebê foi encontrado próximo às margens da Lagoa da Pampulha por um argentino. O bebê estava enrolado em uma manta, e a mãe da criança foi presa horas depois do ocorrido. O americano confirmou em depoimento que a mãe abandonou a criança no local e fugiu logo em seguida.
(Last night, a baby was found close to the margins of the Pampulha Lake by an Argentinian. The baby was wrapped in a blanket, and the baby's mother was arrested hours after the event. The American confirmed in his testimony that the mother abandoned the child at the site and fled right after that.)

² A Latin square distribution, with the reading of the same texts, with or without errors, was not performed, as this would require a higher number of participants, or even a higher number of experimental texts to be read by session, which would render the experiment unfeasible, considering the difficulty in the participation of professional revisers, who had to travel to the experiment site, and considering how tiresome an excessively long the experiment would become.

The experimental texts and the distracters were distributed into 8 topics. In the construction of experimental texts, variables that can affect the results were controlled: 3 sentences; 54 words per text, 18 per sentence; the target words were the 15th and the 16th in the missing preposition text, and the 38th in the incorrect anaphora texts, while the incorrect referent was the 18th word in the text. In the missing preposition text, the word before the preposition had three syllables, and the word after the preposition had 2 syllables, in the Portuguese version. In the texts with the incorrect anaphora, both the referent and the anaphora had 4 syllables, in the Portuguese version.

Specifically regarding the texts with an incorrect anaphora, the informational structure of the texts was also taken into account: the referent (new lexical item introduced in the first sentence text) is not the subject of the sentence, and is accompanied by an indefinite article. Therefore, it is not a focused word. The incorrect anaphora, in turn, is the subject in the third sentence and is accompanied by a definite article, indicating that it was a reference to an item previously introduced in the discourse. Distracter texts were carefully constructed including sentences with a similar syntactic structure, but without an incorrect anaphora, to minimize the possibilities of the participants anticipating the existence of error due to the syntactic structure.

As the frequency of use is quite often related to the fixation time, the Portuguese Bank (PB) corpora, compiled by Tony Sardinha at the Catholic University of São Paulo, was used to control this variable. In general, experimental words are more common in the missing preposition texts than they are in incorrect anaphora texts.

The free program “Coh-Metrix-Port”³, which analyzes several parameters, was used, aimed particularly at offering subsidies to the text readability analysis. The Flesch index is a superficial measure of a text’s readability, which takes into account the average sizes of words and sentences, relating them to the ease of reading. The higher the value obtained, the easier the text processing. Considering this index, the texts in the experiment are classified as very easy, easy or difficult, that is,

³ Available at <www.nilc.icmc.usp.br/coh-metrix-port>. The Coh-Metrix tool was developed at the University of Memphis and calculates text coherence, using different measurements. In Brazil, the Portuguese adaptation was done by researchers at the Interinstitutional Linguistic Center at the University of São Paulo at São Carlos.

corresponding to the grades before higher education. Taking this index into account, as the subjects that participated in the research are at least college students, they should have no problems reading the text.

Since the predictability of words may affect the time focused on them, the predictability of the target word in experimental texts was controlled, by means of a completion test (frame), performed with students majoring in Languages at the Federal University of Minas Gerais (UFMG). For the missing preposition cases, the target word, that is, the one after the preposition, is predictable, since the correct word was used to complete the sentence between 53% and 100% of the times. For incorrect anaphora texts, the target words (the referent and the anaphora) are not predictable, as the chosen words were indicated between 0% and 28% of the times.

Finally, before the start of data collection, a pre-test was conducted with 16 students in the graduate Text Revision program at the Pontific Catholic University of Minas Gerais (PUC-Minas), aimed at checking which errors/inappropriacies would be detected in both experimental and distracter texts. Thus, it was possible to make the last adaptations to the experimental texts, so as to minimize the possibility of the subjects reaching unexpected detections.

4.3 Performing the task

To investigate the reading by professional text revisers, an error detection experiment was performed, using journalistic texts, in which participants were supposed to read the texts provided on a computer screen, clicking on the errors/inappropriacies found using the mouse. The eye tracker, produced by SR Research, model EyeLink 1000, available at the Psycholinguistics Laboratory at UFMG, was used. This is a desktop tracker, that is, the camera and the lights are installed under the computer screen where the stimuli are projected, 40 to 70 cm away from the subject's head. A head stabilizer was used to minimize the subjects' movements and to allow for the recording of the pupillary and corneal reflex. Although the tracker model used allows for binocular recording, only one of the subject's eyes was recorded, which was defined before the start of data collection, allowing for better equipment calibration.

The experiment was conducted at the Psycholinguistics Laboratory at the UFMG School of Languages. The texts were presented on a computer screen, in a random order. First, the instructions were

provided, the equipment was calibrated and the participants underwent a small reading and error detection training using three short texts. Participants were asked to read the text and try to identify the errors and inappropriacies and, when they identified an error/inappropriacy, they should click on them using the mouse. In addition to the eye movement, a chronometric measure was also obtained: the response time. Furthermore, the instructions also stated that the texts were constructed specifically for the research and, therefore, the facts narrated were not necessarily true.

If any participant requested detailed information about the revision activity, the researcher only informed the participant that he/she was expected to behave as if he/she were revising the text, detecting errors and inappropriacies. If the participants asked about the type of errors to be detected, the researcher informed them that he/she could not answer that, emphasizing that they should mark any errors found. Such procedure was adopted in order to determine how the subjects viewed the revision task, based on the types of error detected.

The revision of each text should not take longer than 1 minute/60 seconds, since, if it exceeded this time limit, the program would forward automatically to the subsequent text. A drift correction screen, on which a circle similar to that of calibration was projected at the exact location where the new text would be shown, was placed between texts. As the subject was supposed to focus on this circle so that the researcher could move on to the next text, it was guaranteed that, at the beginning of the reading, the subject would focus his/her gaze on the first word of each text. This avoids variations among texts and subjects regarding when they start reading each text. The background was white and the letters were black, in Times New Roman, size 20, spaced at 3.5. All efforts were made to remove the computer screen brightness, to avoid visual fatigue.

4.4 Data analysis

Different variables were analyzed, at several text levels: sentence, segment, and target. Therefore, the data analysis was based on two angles: the level at which the investigated variable is applied (text, sentence, or segment), and the relation between eye movement and error detection by the subjects.

The variable investigation level has a direct relationship with the type of error investigated: the missing preposition, as it is a local scope error, requires measures that apply to the level of the word or the

phrase, whereas the scope of the incorrect anaphora lies in the integration of parts of the text and, thus, requires measures that correspond to the sentence and text levels.

Furthermore, this work includes both eye movement measurements and the chronometric measurement of the response time. Thus, it was possible to investigate the eye movement as well as the relation between the eye movement and detection. Therefore, the choice was made to use traditional measurements in reading and eye movement studies, as well as to adapt some measurements to this research, combining them with the response time.

Chart 3 – Dependent variables

	Text	Sentences	Segment
Eye movement	<ul style="list-style-type: none"> – Total reading time – Number of fixations on the texts – Average duration of fixations on the texts 	<ul style="list-style-type: none"> – First-pass fixation time on target sentences – Total fixation times on target sentences – Second-pass fixation time on target sentences – Regression-path reading time on target sentences 	<ul style="list-style-type: none"> – First-pass fixation time – Total fixation times on target segment – Number of fixations on the target segment – Regression-path
Eye movement and detection	<ul style="list-style-type: none"> – Total reading time up to detection – Total reading time after detection – Number of fixations up to detection – Number of fixations after detection – Average duration of fixations up to detection – Average duration of fixations after detection 	<ul style="list-style-type: none"> – Total fixation time on target sentences up to detection – Total fixation time on target sentences after detection – Second-pass fixation time in target sentences up to detection – Second-pass fixation time in target sentences after detection 	<ul style="list-style-type: none"> – Total fixation time on the target segment up to detection – Total fixation time on the target segment after detection

In the measurements grouped as “eye movement”, both the texts in which the errors were detected and those in which they were not detected, were taken into account, in order to compare detection to non-detection. In the measurements grouped as “eye movement and detection”, only texts in which the error was detected, as the time of detection (when the participant clicked on the error using the mouse) was used as a divider to analyze each measurement up to and after such a detection: for example, the total reading time up to the error detection and the total reading time after the error detection, measurements that correspond to the total reading time up to the moment when the participant clicks on the error using the mouse and to the total reading time after this moment, until the task is completed.

The total reading time corresponds to the total time the subject took to read and revise each text. The total reading time up to error detection corresponds to the time the subject took to read the text until clicking on the error using the mouse; and the total reading time after detection⁴ corresponds to the time the subject took until he clicked on the error using the mouse until the text revision is completed. The same division applies to the number of fixations on the text up to and after error detection.

The average duration of fixations is equivalent to the ratio between the total reading time, including fixations and saccades, and the number of fixations on the text. Therefore, the average duration of fixations up to error detection corresponds to the ratio between the total reading time up to detection and the number of fixations up to detection. The same logic applies to the average duration of fixations after the error detection.

The first-pass fixation time corresponds to the sum of durations of all fixations performed (whether on the target sentence, at the sentence level, or on the target word, at the local level) until it is abandoned to the left or to the right, as well as during the first reading, that is, it is a measure that indicates the time required when first reading the target.

The total time of fixation on the sentences, however, is equivalent to the sum of all fixations performed on the sentence. This time was also

⁴ Although a significant difference between experimental condition as to the time measurements after error detection, due to the location of each target section (on the first sentence in the case of suppression of the preposition, and on the third sentence for incorrect anaphora), such measurements have been investigated to analyze whether the groups would reread the texts and to verify their engagement in the revision task.

divided into two parts: up to and after error detection, following the same accounting logic described above.

The second-pass fixation time regards the difference between total time of fixation on the sentence and the first-pass fixation time, that is, it is equivalent to the time for return to the sentence after it had been read for the first time. For this measurement, the value up to and after the error detection was also counted.

The last measurement at the sentence level, the regression-path reading time, corresponds to the sum of all fixations performed by the subject from the moment when he/she fixates the target sentence for the first time until the moment he/she fixates on a subsequent part of the text, counting, in addition to the fixations in the target sentences themselves, the possible regressions to previous parts of the text.

The total time of fixation on the target segment is related to the sum of all fixations on it, and the sum of all fixations up to the error detection and after detection was also checked.

The regression-path corresponds to the sum of all fixations performed by the subject from the time he/she fixates the target segment for the first time up to the moment he/she fixates on a subsequent section of the text, counting, in addition to the fixations on the target segment itself, the possible regressions to previous parts of the text.

The independent variables in the eye movement analysis for experimental texts were:

1. Types of errors:

- a. missing preposition
- b. incorrect anaphora

2. Text revision groups:

- a. revisers
- b. non-revisers

3. Detection:

- a. errors detected
- b. errors not detected

The crossing of these variables generates 8 groups:

1. revisers – missing preposition – error detected;
2. non-revisers – missing preposition – error detected;
3. revisers – missing preposition – error not detected;
4. non-revisers – missing preposition – error not detected;
5. revisers – incorrect anaphora – error detected;
6. non-revisers – incorrect anaphora – error detected;
7. revisers – incorrect anaphora – error not detected;
8. non-revisers – incorrect anaphora – error not detected;

First, the data were tabulated and explored descriptively, checking the central tendency measures, such as mean, median, and standard deviation, as well as the visual representation by means of charts, such as boxplot, bar chart, density, and average charts, selected according to the type of variable investigated.

The inferential statistical analysis was then performed, by choosing the most appropriate statistical test based on the analyzed data. The significance level adopted was 95%, therefore, α was defined as 0.05. The program used for statistical analysis was R⁵.

In the variables where the samples adopted a normal distribution, the parametric test was selected. As the experiment was performed with repeated measurements, the defined parametric test was the mixed linear regression model⁶.

In the variables where the samples did not adopt a normal distribution, however, non-parametric tests were selected, namely the Kruskal-Wallis or the Wilcoxon-Mann-Whitney tests, given that these are rank tests and, therefore, knowledge of the sample distribution is required.

5. Results

The revisers detected 191 of the 279 errors, which correspond to 68.5% of the errors. The non-revisers, however, detected 165 of the

⁵ Available at <<https://www.r-project.org/>>.

⁶ This test was also selected, because it enables one to analyze the relation between the variables throughout the task execution, indicating important aspects of the participants' behaviors. This information is not explored in this article, but it is available in Leite (2014).

280 errors, which correspond to 58.9% of the errors. Therefore, revisers were more proficient and detected more errors than did the non-revisers; however, considering only the experimental texts, this difference proved to be statistically insignificant.

The tables provided in this study specify the median values or the “t” or “p” value (depending on the statistical test). Leite (2014) provides more details, such as the number of values considered by variables, the standard deviations, the degrees of freedom and charts. The codes used in all tables provided are as follows:

rev: revisers

nrev: non-revisers

det: error detected

ndet: error not detected

mp: missing preposition

an: incorrect anaphora

N/A: not applicable

* : statistically significant difference

. : marginally significant difference

5.1 Text level

The variables in which statistically significant results were detected are presented below.

In the variables in which the measurements up to and after the error detection were considered, statistically significant results found between the types of errors, that is, when an incorrect anaphora was compared to a missing preposition, were not presented, since these differences were expected due to the composition of texts (in the first case, the error is found in the third sentence, and, in the second text, it is found in the first sentence). Nevertheless, in these cases the revision groups with the same type of error (revisers in anaphora vs. non-revisers in anaphora, and revisers in the missing preposition vs. non-revisers in the missing preposition) were compared.

Table 1 – Summary of the average values of the results obtained in variables at the text level – mixed linear regression models

Measurement	Average		t value	Average		t value	Average		t value
	rev	nrev		an	mp		det	ndet	
total text reading time	35473 ms	27402 ms	*rev>nrev 0.002	32709 ms	30147 ms	*an>sp 0.04	31845 ms	30697 ms	*det>ndet 0.01
number of fixations on the text	134	111	rev>nrev 0.06	129	117	*an>sp 0.002	123	123	-
average duration of fixations on the text	225 ms	205 ms	rev>nrev 0.06	212 ms	218 ms	*sp>an 0.000	218 ms	210 ms	det>ndet 0.05
total reading time after error detection	21590 ms	13812 ms	*rev>nrev 0.001	13323 ms	20696 ms	*sp>an 0.000	N/A	N/A	N/A
average duration of fixations up to error detection	245 ms	222 ms	rev>nrev 0.06	219 ms	243 ms	*sp>an 0.000	N/A	N/A	N/A

Table 2 – Summary of the interaction results obtained in the variables at the text level – mixed linear regression models

Measurement	Variable 1	Variable 2	Average (ms)	Interactions (p value)
Number of fixations in the text	det	mp	104	* an det > mp det (0.003) * an det > an ndet (0.01)
		an	130	
	ndet	mp	113	
		an	112	
Average duration of fixations on the text	det	mp	214	* mp det > an det (0.000) * mp det > an ndet (0.03) * mp det > sp ndet (0.000)
		an	210	
	ndet	mp	198	
		an	207	

Table 3 – Summary of the results obtained at the text level – multiple comparisons of the Wilcoxon-Mann-Whitney test with Bonferroni correction

Measurement	Variable 1	Variable 2	Median	Interactions (p value)
total reading time up to error detection	mp	rev	7469 ms	* rev an > nrev an (0.000)
		nrev	7645 ms	
	an	rev	22301 ms	
		nrev	17094 ms	
number of fixations up to error detection	mp	rev	29	* rev an > nrev an (0.005)
		nrev	34	* nrev mp > rev mp (0.01)
	an	rev	104	
		nrev	81	
average duration of fixations after error detection	mp	rev	276 ms	* rev mp > nrev mp (0.02)
		nrev	265 ms	
	an	rev	377 ms	
		nrev	377 ms	
number of fixations after error detection	mp	rev	81.5	* rev an > nrev an (0.002)
		nrev	56.5	* rev mp > nrev mp (0.000)
	an	rev	38	
		nrev	20	

In comparing the two types of errors, it was determined that the subjects had a longer reading time when the incorrect anaphora was present, which is associated with the number of fixations on the texts, since, also regarding this variable, there was a statistically significant difference between the two conditions, with a higher number of fixations occurring in the texts in which the incorrect anaphora was present. Contrary to this, fixations were longer on average in the texts where the missing preposition was present, which may be also be confirmed by the average duration of fixations up to error detection.

Furthermore, in the interaction among these variables (table 3), the results indicate a difference between the incorrect anaphora detection and the missing preposition detection: while in the incorrect anaphora the higher total reading time is related to the higher number of fixations on the text, in the missing preposition, the detection is related to longer fixations on average. This difference in pattern, therefore, may be related to two different types of reading: a reading towards the surface, with longer fixations, and a reading towards more global levels, with a higher number of fixations on the text.

Regarding the analysis obtained at the text level, the results indicate that revisers, in general, take longer to read the text, and their reading is associated with a higher number of fixations.

In comparing revisers and non-revisers, the first produced higher values for the total reading time and total reading time after error detection (table 1), which indicates that, in addition to taking longer to read the texts, they also maintained the reading longer even after they had detected the error. As to the total reading time up to error detection, in the comparison of revision groups in the same error type, revisers presented higher values than those for non-revisers when the incorrect anaphora was present, but did not exhibit different times regarding the detection of the missing preposition (table 3).

Although there has been a merely marginally significant difference between revisers and non-revisers in terms of the number of fixations on the texts (table 1), upon analyzing this number up to the error detection, in cases of the missing preposition, the number of fixations was higher for the group of non-revisers, whereas in incorrect anaphora it was higher, on average, for the group of revisers. After the error detection, the revisers performed more fixations, regardless of the type of error.

As for the average duration of fixations on the text, there was a marginally significant difference between how the revisers and the non-revisers performed, as also occurred when this variable was analyzed up to the error detection (table 1), with revisers tending to perform longer fixations, on average. There was a statistically significant difference between the groups regarding the average duration of fixations after error detection (table 3).

Therefore, revisers showed a different reading pattern for each type of error: in detecting the missing preposition, the fixations tended to be longer (marginally significant average duration of fixations), and in detecting the incorrect anaphora, the number of fixations was higher.

The total reading time was also longer in the texts in which the error was detected than in those texts in which the subject did not detect the error, and the number of fixations was also higher for those texts; however, this finding may result from the task itself, as, in the texts in which the error was detected, the participant took some time clicking on the error using the mouse.

5.2 Sentence level

Regarding the measurements presented at the sentence level, in the missing preposition cases, only the first sentence, in which the error was located, was taken into account. For incorrect anaphora cases, both the first and the third sentences were analyzed, as they contained the referent and the incorrect anaphora, respectively.

Table 4 – Summary of results obtained from variables at the sentence level – mixed linear regression models

Measurement	Average (ms)		t value	Average (ms)		t value	Average (ms)		t value
	rev	nrev		an	mp		det	ndet	
total fixation time on the first sentence	11274	9227	*rev>nrev 0.02	9971 ms	10527	0.45	10592	9646	*det>ndet 0.000
total fixation time on the third sentence	11050	8240	*rev>nrev 0.009	N/A	N/A	N/A	10756	8655	*det>ndet 0.000
total fixation time on the first sentence up to error detection	9075	8270	*rev>nrev 0.008	9378	8305	0.24	N/A	N/A	N/A
total fixation time on the third sentence up to error detection	6341	4868	0.21	N/A	N/A	N/A	N/A	N/A	N/A

Table 5 – Summary of interaction results obtained from variables at the sentence level – mixed linear regression models

Measurement	Variable 1	Variable 2	Average (ms)	Interactions (p value)
total fixation time on the first sentence up to error detection	mp	rev	8192	* rev an > nrev an (0.03) * rev an > rev sp (0.03)
		nrev	8431	
	an	rev	10522	
		nrev	7981	

Table 6 – Summary of the results obtained at the sentence level – multiple comparisons of the Wilcoxon-Mann-Whitney test with Bonferroni correction

Measurement	Variable 1	Variable 2	Variable 3	Median (ms)	Interactions (p value)
second-pass fixation time on the first sentence	det	mp	rev	2241	* rev an det > nrev an det (0.000) * rev an det > nrev mp det (0.000) * rev an det > rev mp det (0.000) * rev an det > nrev an ndet (0.000) * rev an det > rev an ndet (0.000) * rev an det > nrev mp ndet (0.000) * rev mp det < nrev mp det (0.000) * nrev mp det < nrev an ndet (0.03) * nrev mp det < rev an ndet (0.03) * nrev mp det < rev mp ndet (0.01)
			nrev	182.5	
		an	rev	5076	
			nrev	2088	
	ndet	mp	rev	4274	
			nrev	1846	
		an	rev	1668.5	
			nrev	2092	
first-pass fixation time on the third sentence	an	det	rev	5338	* nrev det > nrev ndet (0.000) * nrev det > rev ndet (0.000) * rev det < nrev ndet (0.006)
			nrev	6181	
		ndet	rev	6072	
			nrev	5739	
second-pass fixation time on the third sentence	an	det	rev	8659	* rev det > nrev det (0.000) * rev det > nrev ndet (0.000) * rev det > rev ndet (0.000) * nrev det > nrev ndet (0.000) * nrev det > rev ndet (0.01)
			nrev	6039	
		ndet	rev	2118.5	
			nrev	1666	
regression-path reading time on the third sentence	an	det	rev	4223.5	* rev det > nrev ndet (0.000) * rev det > rev ndet (0.000) * nrev det > rev ndet (0.002) * nrev det > nrev ndet (0.000)
			nrev	4282	
		ndet	rev	2489.5	
			nrev	2199	
total fixation time on the first sentence after error detection	det	mp	rev	2208.5	* rev an > nrev an (0.001) * rev an < rev sp (0.000) * nrev an < nrev mp (0.000) * nrev an < rev mp (0.000) * rev sp > nrev mp (0.000)
			nrev	525.5	
		an	rev	506.0	
			nrev	0	
total fixation time on the third sentence after error detection	det	an	rev	5232	* rev > nrev (0.000)
			nrev	2933	
second-pass fixation time on the first sentence up to error detection	det	mp	rev	0	* rev an > nrev an (0.02) * rev an > rev mp (0.000) * rev an > nrev mp (0.000) * nrev an > rev mp (0.000) * nrev an > nrev mp (0.000)
			nrev	0	
		an	rev	2904.5	
			nrev	1526	

When the performances of revisers and non-revisers were compared, there was a statistically significant difference as to the total fixation time on the first and third sentences of the text, as to the total fixation time on the first sentence of the text up to the error detection, as well as to the total fixation time on the first and third sentences after error detection. In all of these situations, whether in the cases of incorrect anaphora or the missing preposition, the revisers presented higher values on average than did those obtained for non-revisers. The two groups were also different as to the second-pass fixation time on the first and third sentences, to which revisers took longer to return than did the non-revisers. Revisers also took longer to return to the first sentence up to the detection of the incorrect anaphora (second-pass fixation time up to the error detection) than did the non-revisers. However, neither group presented different results regarding the missing preposition, since, for both groups, the detection of the missing preposition took place on the first reading of the sentence.

5.3 Local level

Within the condition of the missing preposition, only the total fixation time on the target segment, including the 15th and the 16th words of the text, were analyzed. For the condition of incorrect anaphora, however, the total fixation times on the referent and on the anaphora were analyzed separately.

Table 7 – Summary of results obtained from variables at the local level – mixed linear regression models

Measurement	Average		t value	Average		t value	Average		t value
	rev	nrev		an	mp		det	ndet	
number of fixations on the target – mp	10.74	11.74	0.16	N/A	N/A	N/A	11.91	7.68	*det>ndet 0.000
total fixation time in the referent up to error detection	1169 ms	831 ms	*rev>nrev 0.007	N/A	N/A	N/A	N/A	N/A	N/A

Table 8 – Summary of interaction results obtained from variables at the local level – mixed linear regression models

Measurement	Variable 1	Variable 2	Average	Interactions (p value)
number of fixations on the target – mp	det	rev	11	* rev det > rev ndet (0.000) * rev det > nrev ndet (0.000) * nrev det > rev ndet (0.000) * nrev det > rev ndet (0.000)
		nrev	13	
	ndet	rev	8	
		nrev	7	

Table 9 – Summary of the results obtained at the local level – multiple comparisons of the Wilcoxon-Mann-Whitney test with Bonferroni correction

Measurement	Variable 1	Variable 2	Median	Interactions (p value)
total fixation time on the target – mp	det	rev	2980 ms	* nrev det > nrev ndet (0.000) * nrev det > rev ndet (0.000) * rev det > nrev ndet (0.000)
		nrev	3233 ms	
	ndet	rev	1072 ms	
		nrev	1085 ms	
total fixation time on the referent	det	rev	1229 ms	* rev det > nrev det (0.002) * nrev det > nrev ndet (0.008) * rev det > nrev ndet (0.000)
		nrev	726 ms	
	ndet	rev	724.5 ms	
		nrev	497 ms	
total fixation time on the anaphora	det	rev	2092 ms	* rev det > nrev ndet (0.000) * rev det > rev ndet (0.000) * nrev det > nrev ndet (0.000)
		nrev	1799 ms	
	ndet	rev	669 ms	
		nrev	475 ms	
number of fixations on the referent	det	rev	6	* nrev det > nrev ndet (0.003) * rev det > nrev ndet (0.000) * rev det > rev ndet (0.000)
		nrev	4	
	ndet	rev	3	
		nrev	3	
number of fixations on the anaphora	det	rev	8	* nrev det > nrev ndet (0.000) * rev det > nrev ndet (0.000) * rev det > rev ndet (0.000)
		nrev	7	
	ndet	rev	3	
		nrev	2	
total fixation time on the target after error detection – mp	det	rev	675.5 ms	* rev > nrev (0.000)
		nrev	104 ms	
total fixation time on the anaphora after error detection	det	rev	246.5 ms	* rev > nrev (0.000)
		nrev	0 ms	

When the total fixation time on the target segment during the entire text reading was detected, not taking into account the detection time, no statistically significant differences were found between the performances of revisers and non-revisers. Actually, there were statistically significant differences between the texts in which the errors were detected, and those in which the errors were not detected, and, whenever the detection occurred; the total fixation time on the target was longer. This pattern was found in texts with cases of the missing preposition, as well as in those with cases of incorrect anaphora. For this type of error, the difference between the performances of revisers and non-revisers was only identified on the referent, as values were higher for the first. The same results were obtained in the analysis of the number of fixations on the target, but, in this case, the difference between revisers and non-revisers regarding the number of fixations on the referent and on the anaphora was only marginally significant.

Similarly, upon analyzing the total fixation time on the target up to the error detection, that is, the sum of all fixations performed on the target until the participant clicked on the mouse, including the regressions to the word up to that point, a significant difference between the performances of revisers and non-revisers was detected only regarding the referent, given that, in incorrect anaphora cases, the values in general in this case were higher for the performances of revisers. In analyzing the total fixation time on the referent after error detection, that is, after the participants had clicked on the error using the mouse, however, no significant differences between the performances of revisers and non-revisers were detected, as both groups, most of the time, did not return to the referent.

Regarding the incorrect anaphora area, a statistically significant difference was detected between the performances of revisers and non-revisers, as to the total fixation time after the error detection. Therefore, the revisers spent more time fixating the referent to detect the incongruity and possibly attempt to solve it. After the error detection, they returned more times to the anaphora before completing the text reading.

In the missing preposition cases, similarly to what occurred to the total fixation time on the target, no statistically significant differences were found between the performances of revisers and non-revisers as regards the total fixation time on the target up to error detection. The opposite occurred as to the total time of fixation on the target after error detection: revisers obtained higher values than those obtained by the non-revisers, on average.

Therefore, although no significant difference between the performances of revisers and non-revisers was detected when the total fixation time on the target in the text reading was analyzed, this difference did occur when the eye movement was analyzed, taking into account the error detection time. In this case, revisers, in general, obtained higher values than those obtained by non-revisers and, in the texts in which the errors were detected, the total fixation times on the target were higher than in those texts in which the errors were not detected. This variable, therefore, is relevant for the investigation of proficiency in error detection.

6 Final considerations

The results indicated that, in general, the texts in which the incorrect anaphora was present yielded higher values in the investigated measurements, especially at the text level, and even had a longer reading time, related to a higher number of fixations. This result corroborates prior studies in this area, indicating that the processing of errors at levels that require the integration of parts of the text is more difficult. The exception was the average duration of the fixations on the text, as well as the error detection. In this case, the values were higher for texts in which the preposition was suppressed. Since in this measure the interaction between the error type and the detection was significant, it may be stated that the increase in the average duration of the fixations on the text, including in the start, up to error detection, favors the identification of the missing preposition. Considering the location of the errors on the text and their complexity, these differences may result from the global strategies used in the reading at an early or late stage. Hence, it is important that other studies explore this initial finding, including the comparison of a base condition, with and without errors.

The results of this research corroborate Vauras, Hyona, and Niemi's proposal (1992), according to whom incoherences increase both the fixation time and the number of regressions. This research, therefore, confirmed Linguistics data that set the difference between readings directed to levels that are more global and readings directed to levels that are more superficial, also contributing to the characterization of these types of reading, as well as to showing which strategies may be more appropriate for each level. In this sense, an important, though initial, finding is that the detection of the missing preposition is, on

average, favored when longer fixations are performed. Many studies in the area proposed that the reading directed to superficial levels must be different from reading directed to more global levels, although they do not specify how this difference is processed regarding eye movement. The longer duration of fixations appears to be a significant parameter, though other studies are required to confirm this finding, including those with different types or surface errors.

Regarding the differences between those texts in which the error was detected and those in which the error was not detected, the general perspective is that error detection is followed by increased fixation time, at the text, sentence, or local level. At the local level, in the texts in which the error is detected, the total fixation time on the target is higher, which occurs both for the missing preposition and in incorrect anaphora. Such a difference, however, may result from the time required by the participants to click on the text using the mouse.

Finally, regarding the difference between the performances of revisers and non-revisers, the general perspective was that revisers presented higher values for the variables in which a significant difference occurred, which means they were slower in reading than the non-revisers. This occurred at the text, sentence, and local levels.

Thus, this study's results confirmed the initial hypothesis that professional revisers would require a lengthier, less automatic, more detailed, and controlled reading. According to Hayes (2004), with extensive practice, certain edition aspects may become automatic, but not the edition activity as a whole, even for very experienced subjects, a reason why the reading control is an assumption of revision. When Klein and Hoffman (1992) discuss expertise, they emphasize that the differentiated performance of an expert may be identified based on several factors: variability-consistency, accuracy, completeness, and speed. As to speed, the authors highlight that, in some cases, experts may take longer than novices do to complete a task, even when the novices adopt a very fast behavior, on impulse. This research indicates the occurrence of this type of situation.

Professional revisers were also more specific as to the regressions performed. In texts that presented incorrect anaphora, revisers spent a longer total fixation time on the first sentence up to error detection than did the non-revisers, as well as a longer total fixation time on the referent. Since subjects, in general, chose to return to previous parts of the text, when they encountered the incorrect anaphora, the revisers were more

specific and fixated, for a longer period of time, on the part of the text that was actually important for problem-solving, that is, the referent and the sentence containing it. Furthermore, after detecting the incorrect anaphora, the revisers made longer returns to the third sentence of the text, as well as the anaphora itself, which also indicates that they were more specific in this case. According to Hyona and Nurminen (2006), proficient readers tend to direct their regressions to actually informative parts of the text.

Finally, it should be considered that this research specifically investigated error detection, which is one of the revision stages. In addition to this, the corrections must also be performed, but this stage was not investigated by this study, though it deserves to be considered in further studies in the area, including those in which the quality of corrections performed is determined, hence a qualitative investigation. Considering that studies on professional text revision are scarce, especially those with a cognitive approach, this research had an exploratory nature and provided important initial considerations, not only to the study of performances of professional revisers, but also for the study of text revision and reading processes.

References

ALAMARGOT, D. *et al.* Eye and Pen: a new device to study reading during writing. *Behaviour Research Methods, Instruments and Computers*, Bethesda, USA, v. 38, n. 2, p. 287-299, 2006. doi.org/10.3758/BF03192780

HAYES J. R. A new framework for understanding cognition and affect in writing. In: C. Levy; S. Ransdell (Ed.). *The science of writing: theories, methods, individual differences and applications*. Mahwah: Lawrence Erlbaum Associates, 1996. p. 1-28.

HAYES, J. R. What triggers revision? In: ALLAL, L.; CHANQUOY, L.; LARGY, P. (Ed.). *Revision: cognitive and instructional processes*. Kluwer Academic Publishers, Massachusetts, 2004, p. 9-22. doi.org/10.1007/978-94-007-1048-1_2

HAYES J. R. *et al.* Cognitive processes in revision. In: ROSENBERG, S. (Ed.). *Advances in psycholinguistics: Cambridge: Cambridge University Press*, 1987. v. II: Reading, writing, and language processing, p. 176-240.

HAYES, J. R.; FLOWER, L. S. Identifying the organization of writing processes. In: GREGG, L. W.; STEINBERG, E. R. (Ed.). *Cognitive processes in writing: An interdisciplinary approach*. Lawrence Erlbaum Associates, Hillsdale, 1980. p. 3-30.

HEURLEY, L. La révisión de texto: L'approche de la Psychologie Cognitive. *Langages*, Larousse, v. 164, n. 4, p. 10-25, 2006. doi.org/10.3917/lang.164.0010.

HYONA, Jukka; KAAKINEN, Johanna K.; LORCH, Robert F. Jr. Individual Differences in Reading to Summarize Expository Text: Evidence From Eye Fixation Patterns. *Journal of Educational Psychology*, Taylor & Francis, v. 94, n. 1, p. 44-55, 2002. doi.org/10.1037/0022-0663.94.1.44.

HYONA, Jukka; NIEMI, Pekka. Eye Movements During Repeated Reading of a Text. *Acta Psychologica*, Elsevier, v. 73, p. 259-280, 1990. doi.org/10.1016/0001-6918(90)90026-C.

HYONA, Jukka; NURMINEN, Anna-Mari. Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology*, John Wiley & Sons, v. 97, p. 31-50, 2006. doi.org/10.1348/000712605X53678.

KLEIN, G. A.; HOFFMAN, R. R. Seeing the invisible: perceptual-cognitive aspects of expertise. In: RABINOWITZ, M. (Ed.). *Cognitive Science Foundations of Instruction*. Mahwah, N. J: Erlbaum, 1992. p. 203-226.

LAND, Michael F. Fixation Strategies during active behavior: a brief history. In: VAN GOMPEL, R. P. G. et al. (Ed.). *Eye Movements: a window on mind and brain*. Amsterdã: Elsevier, 2007. p. 76-95.

LEITE, Délia Ribeiro. *O olhar do profissional: estudo do movimento ocular na leitura realizada por revisores de textos*. 2014. Tese (Doutorado) – Universidade Federal de Minas Gerais, Belo Horizonte, 2014.

LUEGI, Paula. *O registro do movimento dos olhos durante a leitura de textos*. Dissertação de mestrado. Universidade de Lisboa, 2006.

RAYNER, Keith. Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, Washington, v. 124, n. 3, p. 372-422, 1998. doi.org/10.1037/0033-2909.124.3.372.

STAUB, Adrian; RAYNER, Keith. Eye movements and on-line comprehension processes. In: GASKELL, G. (Ed.). *The Oxford Handbook of Psycholinguistics*. Oxford, UK: Oxford University Press, 2007. p. 327-342.

VAURAS, Marja; HYONA, Jukka; NIEMI, Pekka. Comprehending coherent and incoherent texts: evidence from eye movement patterns and recall performance. *Journal of Research in Reading*, Wiley online Library, v. 15, n. 1, p. 39-54, 1992. doi.org/10.1111/j.1467-9817.1992.tb00020.x.

WAGNER, Luiz Roberto; CUNHA, Djenane S. W. *Categorias de um revisor*. Língua Portuguesa – conhecimento prático, 2012. n. 37, p. 12-13.