

REVISTA DE ESTUDOS DA LINGUAGEM

Universidade Federal de Minas Gerais

REITORA: Ana Lúcia Almeida Gazzola

VICE-REITOR: Marcos Borato Viana

Faculdade de Letras

DIRETORA: Eliana Amarante de Mendonça Mendes

VICE-DIRETORA: Veronika Benn-Ibler

Comissão Editorial

Marco Antônio de Oliveira (Presidente)

Vera Lúcia Menezes de Oliveira e Paiva

Maria Antonieta Amarante M. Cohen

João Carlos de Melo Mota

Luiz Carlos de Assis Rocha

Carlos Alberto Marques dos Reis

Conselho Editorial

Adriana Silvina Pagano (UFMG)

Anthony Julius Naro (UFRJ)

Ataliba T. Castilho (USP)

Beatriz Vaz Leão (UFMG)

Carlos Alberto Gohn (UFMG)

Carmen Rosa Caldas-Coulthard

(The University of Birmingham)

Charlotte C. Galves (UNICAMP)

Clara Grimaldi Eleazaro (UFMG)

Demerval da Hora (UFPB)

Francis Aubert Lourado (PUC/SP)

Gregory R. Guy (York University)

Heliana Ribeiro de Mello (UFMG)

Hilário I. Bohn (UFSC)

Hildo Honório do Couto (UNB)

Hugo Mari (UFMG)

Ida Lúcia Machado (UFMG)

Johnny José Mafra (UFMG)

José Luiz Fiorin (USP)

Kanavillil Rajagopalan (UNICAMP)

Leda Bisol (PUC/RS)

Leila Bárbara (PUC/SP)

Leo Wetzels (Free Univ. of Amsterdam)

Lorenzo Vitral (UFMG)

Luiz Antônio Marcushi (UFPE)

Luiz Carlos de Assis Rocha (UFMG)

Luiz Carlos Cagliari (UNICAMP)

Luiz Carlos Travaglia (UFU)

Marco Antônio de Oliveira (UFMG)

Marco Antônio Rodrigues Vieira (UFV)

Maria Antonieta Amarante M. Cohen (UFMG)

Maria Bernadete Marques Abaurre (UNICAMP)

Maria Carlota Amaral Paixão Rosa (UFRJ)

Maria Cecília Magalhães Mollica (UFRJ)

Maria Luíza Braga (PUC/RJ)

Marta M. P. Scherre (UnB)

Milton do Nascimento (PUC/MG)

Rodolfo Ilari (UNICAMP)

Seung-Hwa Lee (UFMG)

Sírio Possenti (UNICAMP)

Tânia Maria Alkimim (UNICAMP)

Thaís Cristóvão Silva (UFMG)

Vera Lúcia Menezes de Oliveira e Paiva (UFMG)

Vilson J. Leffa (UFRGS)

Yone de Freitas Leite (Museu Nacional)

Revisão

Soélis Teixeira do Prado Mendes

Revisão final

César Reis

José Olímpio de Magalhães

Capa e projeto gráfico original

Elson Rezende de Melo

Projeto gráfico atualizado e Editoração eletrônica

Marco Antônio & Alda Durães

REVISTA DE ESTUDOS DA LINGUAGEM

V. 12 - Nº 2- JUL./DEZ. 2004

Ficha Catalográfica:

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: 0104-0588

Sumário

Apresentação <i>César Reis</i>	11
---	----

FONÉTICA EXPERIMENTAL

A influência da natureza fonética do fonema inicial da palavra na medida de tempo de reação em tarefas de leitura em voz alta <i>Ângela Maria Vieira Pinheiro</i> <i>César Reis</i>	19
---	----

Analysis of acoustic-prosodic features of spontaneous expressive speech <i>Carlos Toshinori Ishi</i> <i>Nick Campbell</i>	37
--	----

O uso de estímulos sintéticos em testes de percepção de vogais de uma língua estrangeira <i>Ricardo Augusto Hoffmann Bion</i>	51
---	----

Para a melhoria da síntese articulatória das vogais nasais do português europeu: estudo da duração e de características relacionadas com a fonte glotal <i>António Teixeira</i> <i>Francisco Vaz</i> <i>Lurdes Castro Moutinho</i> <i>Rosa Lídia Coimbra</i> <i>Raquel Castro Lisboa</i>	65
---	----

Alguns esquemas entoacionais característicos da leitura de um texto por portugueses e brasileiros <i>Antônio Romano</i> <i>Lurdes de Castro Moutinho</i>	93
Análise comparada de três padrões prosódicos em francês e em português europeu <i>Lurdes de Castro Moutinho</i> <i>Jean-Pierre Zerling</i>	111
Intensidade e duração de formas reduzidas no português brasileiro <i>Larissa Santos Ciríaco</i> <i>Lorenzo Teixeira Vitral</i> <i>César Reis</i>	143
Tarefa de discriminação de fonemas com pseudopalavras <i>Rui Rothe-Neves</i> <i>Regina Carla Lapate</i> <i>Juliana Sardinha Pinto</i>	159
Reorganização prosódica na disartria <i>Erica Reviglio Iliovitz</i>	169
A pausa na produção da fala com comprometimento neurológico <i>Jussara Melo Vieira</i> <i>Plínio Almeida Barbosa</i> <i>Maria Inês Pegoraro-Krook</i>	181
Banco de dados orais: uma nova perspectiva aos estudos do português brasileiro <i>Regina Célia Fernandes Cruz</i> <i>Jailma do Socorro Uchôa Bulhões</i> <i>Léa da Silva Fernandes</i>	193

TEORIA FONOLÓGICA

O caminho para a realização de / l / e / r / intervocálicos na aquisição com atraso: a substituição por [j] <i>José Olímpio de Magalhães</i> <i>Maíra Martins</i>	215
Seqüências de (sibilante + consoante) no português de Belo Horizonte <i>Thais Cristófaró-Silva</i> <i>Daniela Mara Lima Oliveira</i>	247
O alicamento das vogais médias pretônicas e o conceito de léxico com armazenamento exemplar <i>Maria do Carmo Viegas</i>	269
A importância da proeminência da frase fonológica no Português Brasileiro <i>Luciani Tenani</i>	289
Fonologia Prosódica e Teoria da Otimalidade: reflexões sobre a interface sintaxe e fonologia na formação de sintagmas fonológicos <i>Filomena Sandalo</i>	319
Retração acentual e verbos transitivos com leitura intransitiva <i>Raquel S. Santos</i>	345
Da frase fonológica no dialeto goiano <i>Irene Zasiowicz Pinto Calaçã</i>	365
Diminutive Reduplication in Government Phonology <i>Renée Lambert</i>	383
Breve estudo da estrutura silábica do Português arcaico: a posição do glide <i>Fernanda Elias Zucarelli</i>	405

Traço fonológico: variação sob a perspectiva da multilinearidade <i>Teresinha de Moraes Brenner</i>	423
Simbolismo sonoro do PB: o estudo dos Ideofones <i>Regina Célia Fernandes Cruz</i> <i>Helane de Fátima Gomes Fernandes</i>	439
The Production of English Initial /s/ clusters by Portuguese and Spanish EFL Speakers <i>Andreia Schurt Rauber</i> <i>Barbara Oughton Baptista</i>	459
Categorias fonéticas na aprendizagem de língua estrangeira <i>Barbara O. Baptista</i>	475
A aquisição de encontros consonantais do Inglês iniciados por /s/: estudo de caso com criança brasileira <i>Magliane De Marco</i>	491
Interpretação fonológica e realização fonética durante o aprendizado de uma LE <i>Nina A. Lyubimova</i>	505
A identificação fonêmica do discurso de estudantes brasileiros na língua russa <i>Tanira Castro</i>	515

AGRADECIMENTOS

Larissa Ciríaco, como estagiária do Laboratório de Fonética, nos auxiliou ao longo da edição desta publicação, constituindo-se num elo de ligação importante entre o organizador da revista, os autores e os funcionários responsáveis pela Revista de Estudos da Linguagem.

José Olímpio de Magalhães, Marco Antônio de Oliveira, Seung-Hua Lee, Thaïs Cristófaros-Silva que participaram da Comissão Editorial deste número especial da Revista de Estudos da Linguagem.

Marco Antônio de Oliveira, presidente da Revista de Estudos da Linguagem, Carlos Reis, funcionário responsável pela parte operacional e Alda Lopes, diagramadora.

Cyril Deniaud, do *Laboratoire Parole et Langage* – CNRS, Aix-en-Provence, deu suporte técnico para a última revisão.

APRESENTAÇÃO

Inicialmente, gostaria de agradecer a todos aqueles que apresentaram seus textos para publicação neste número da Revista de Estudos da Linguagem, dedicado aos trabalhos apresentados no VII Congresso Nacional de Fonética e Fonologia e I Congresso Internacional de Fonética e Fonologia, que se realizou em Belo Horizonte em outubro de 2002. Agradeço a paciência com que aguardaram a longa edição da revista. Fizemos o possível para que todos os problemas apresentados tivessem uma solução. Apresentamos nossas desculpas pelas eventuais falhas, que são de nossa inteira responsabilidade. Nossa meta era a publicação antes do Congresso de São Luís, meta essa, que foi atingida, graças à colaboração de colegas, estudantes, funcionários e dos autores.

Os artigos apresentados, com poucas exceções, estão sendo publicados, oferecendo uma visão dos grupos, indivíduos, temas e metodologias que constituem a pesquisa em Fonética e Fonologia da língua portuguesa no país e no exterior. Como há alguns textos de autores portugueses, respeitamos suas convenções ortográficas, da mesma forma como respeitamos as convenções de transcrição fonética dos diferentes autores.

Teoria Fonológica, Língua Estrangeira, Fala Patológica, Aquisição Fonológica, Prosódia, Fonética Experimental são os temas que agrupam o maior número de trabalhos. Na verdade, podemos dividir os trabalhos em dois grandes grupos: Fonética Experimental e Teoria Fonológica. Tomando-se aqui Fonética Experimental num sentido amplo, isto é, a simples utilização de métodos da física, da fisiologia ou da psicologia no estudo dos sons da fala. Fazemos abaixo uma breve apresentação dos trabalhos aqui publicados, uma vez que estarão sempre precedidos do respectivo *abstract*.

Fonética Experimental

Em ANGELA MARIA VIEIRA PINHEIRO (UFMG) e CÉSAR REIS (UFMG), *A influência da natureza fonética do fonema inicial da palavra na medida de tempo de reação em tarefas de leitura em voz alta*, os autores analisam o efeito da natureza fonética de vários fonemas em posição inicial de palavras regulares de alta frequência no tempo de reação em tarefas de leitura em voz alta;

CARLOS TOSHINORI ISHI (Laboratórios ATR/HIS) e NICK CAMPBELL (Laboratórios ATR/HIS), *Analysis of acoustic-prosodic features of spontaneous expressive speech*, com base em vários parâmetros acústicos, visando à síntese da fala, investigam a participação desses parâmetros acústicos, assim como a existência de correlação entre eles, na manifestação de diferentes emoções e atitudes, com base em dados da fala espontânea do japonês;

RICARDO AUGUSTO HOFFMANN BION (UFSC), *O uso de estímulos sintéticos em testes de percepção de vogais de uma língua estrangeira*, realiza testes de percepção com estudantes brasileiros, sendo os estímulos vogais anteriores do inglês;

Em ANTONIO TEIXEIRA (Universidade de Aveiro), FRANCISCO VAZ (Universidade de Aveiro), LURDES CASTRO MOUTINHO (Universidade de Aveiro), ROSA LÍDIA COIMBRA (Universidade de Aveiro), RAQUEL CASTRO LISBOA (Universidade de Aveiro), *Para a melhoria da síntese articulatória das vogais nasais do Português Europeu: estudo da duração e de características relacionadas com a fonte glotal*, os autores apresentam um estudo acústico e glotográfico de vários parâmetros da fonte glotal;

ANTÓNIO ROMANO (Università di Torino) e LURDES DE CASTRO MOUTINHO (Universidade de Aveiro), *Alguns esquemas entoacionais característicos da leitura de um texto por portugueses e brasileiros*, fazem um estudo comparativo de entonações do português brasileiro e português europeu;

LURDES DE CASTRO MOUTINHO (Universidade de Aveiro) e JEAN-PIERRE ZERLING (Universidade Marc Bloch), *Análise comparada de três padrões prosódicos em Francês e em Português*

Europeu, é um estudo prosódico em que são comparados enunciados declarativo, interrogativo e imperativo do português europeu e do francês;

LARISSA CIRÍACO (UFMG), LORENZO TEIXEIRA VITRAL (UFMG), CÉSAR REIS (UFMG), *Intensidade e duração de formas reduzidas no português brasileiro*, apresentam um estudo fonético acústico de algumas formas reduzidas do português brasileiro em processo de cliticização;

RUI ROTHE-NEVES (UFMG), REGINA CARLA LAPATE (UFMG) e JULIANA SARDINHA PINTO (UFMG), *Tarefa de discriminação de fonemas com pseudopalavras*, apresentam uma tarefa de discriminação de fonemas por pacientes afásicos;

ERICA REVIGLIO ILIOVITZ (UNICAMP), *Reorganização prosódica na disartria*, apresenta um estudo prosódico longitudinal de dois pacientes com disartria pós-traumática;

JUSSARA MELO VIEIRA, PLÍNIO ALMEIDA BARBOSA (UNICAMP), MARIA INÊS PEGORARO-KROOK, *A pausa na produção da fala com comprometimento neurológico*, estudam a pausa em um indivíduo com disartria e incompetência velofaríngea, examinando a correlação entre a duração das pausas e a natureza da fronteira sintática;

REGINA CÉLIA FERNANDES CRUZ (UFPA), JAILMA DO SOCORRO UCHÔA BULHÕES (UFPA) e LÉA DA SILVA FERNANDES (UFPA), *Banco de Dados Oraís: uma nova perspectiva aos estudos do português brasileiro*, apresentam um banco de dados orais destinado aos pesquisadores sociolinguistas e foneticistas.

Teoria Fonológica

JOSÉ OLÍMPIO DE MAGALHÃES (UFMG) e MAÍRA MARTINS (UFMG), *O caminho para a realização de /l/ e /r/ intervocálicos na aquisição com atraso: a substituição por [j]*, verificam algumas hipóteses nas etapas de aquisição dos fonemas /l/ e /r/ em crianças com atraso de linguagem;

THAÏS CRISTÓFARO-SILVA (UFMG) e DANIELA MARA LIMA OLIVEIRA (UFMG), *Seqüências de (sibilante + consoante) no português de Belo Horizonte*, examinam a variação sonora nas seqüências (sibilante + oclusiva alveolar) e (sibilante + africada alveopalatal);

MARIA DO CARMO VIEGAS (UFMG), *O alçamento das vogais médias pretônicas e o conceito de léxico com armazenamento exemplar*, apresenta reflexões teóricas sobre o alçamento da vogal média pretônica;

LUCIANA TENANI (UNESP/SJRP), *A importância da proeminência da frase fonológica no português brasileiro*, estuda a inter-relação entre a estrutura prosódica, no caso a frase fonológica, e os processos fonológicos de degeminação e elisão;

FILOMENA SANDALO (UNICAMP), *Fonologia Prosódica e Teoria da Otimidade: Reflexões sobre a interface sintaxe e fonologia na formação de sintagmas fonológicos*, trata de temas teóricos envolvendo as relações entre sintaxe e prosódia, enfocando-os com base na ocorrência do choque acentual;

RAQUEL S. SANTOS (USP), *Retração acentual e verbos transitivos com leitura intransitiva*, trata da retração acentual vista do ponto de vista das relações entre a fonologia e a sintaxe;

IRENE ZASIMOWICZ PINTO CALAÇA (UFMG), *Da frase fonológica no dialeto goiano*, estuda alguns aspectos prosódicos e fonológicos dentro da frase fonológica;

RENÉE LAMBERT (Université Lumière-Lyon 2), *Diminutive reduplication in Government Phonology*, se baseia em dados de reduplicação em várias línguas para fundamentar suas reflexões teóricas sobre a representação morfológica e fonológica de algumas categorias lexicais;

FERNANDA ELIAS ZUCARELLI (Faculdades Integradas FAFIBE), *Breve estudo da estrutura silábica do português arcaico: a posição do glide*, analisa a posição da semivogal na estrutura silábica no português arcaico;

TERESINHA DE MORAES BRENNER (UFSC), *Traço fonológico: variação sob a perspectiva da multilinearidade*, examina alguns

fenômenos de variação fonológica no português brasileiro com base na teoria da Geometria de Traços;

REGINA CÉLIA FERNANDES CRUZ(UFPA) e HELANE DE FÁTIMA GOMES FERNANDES (UFPA), *Simbolismo sonoro do PB: o estudo dos ideofones*, analisam processos de formação de ideofones no PB;

ANDRÉIA SCHURT RAUBER (UFSC) e BARBARA OUGHTON BAPTISTA (UFSC), *The production of English initial /s/ clusters by Portuguese and Spanish EFL speakers*, tratam da produção de encontros consonantais iniciais (sC(C)) em inglês por falantes do português e do espanhol, procurando extrair resultados teóricos e pedagógicos;

BARBARA O. BAPTISTA (UFSC), *Categorias fonéticas na aprendizagem de língua estrangeira*, examina dois modelos de aprendizagem dos sons de língua estrangeira;

MAGLIANE DE MARCO (Universidade Católica de Pelotas), *A aquisição de encontros consonantais do inglês iniciados por /s/: estudo de caso com criança brasileira*, estuda a aquisição de encontros consonantais do inglês iniciados por /s/ por uma criança brasileira, buscando na teoria fonológica explicações para os fatos observados;

NINA A. LYUBIMOVA (UFRGS), *Interpretação fonológica e realização fonética durante o aprendizado de uma LE*, trata de uma unidade fônica especial que surge no processo de aquisição de uma língua estrangeira e que denomina de diafonema;

TANIRA CASTRO (UFRGS), *A identificação fonêmica do discurso de estudantes brasileiros na língua russa*, apresenta um estudo sobre as dificuldades dos estudantes brasileiros na aquisição da pronúncia da língua russa.

Fonética Experimental

A influência da natureza fonética do fonema inicial da palavra na medida de tempo de reação em tarefas de leitura em voz alta

Ângela Maria Vieira Pinheiro
Universidade Federal de Minas Gerais

César Reis
Universidade Federal de Minas Gerais

Abstract

Experiments of reading aloud isolated words, in which the reaction time (RT) is measured based on the time between the appearance of the stimulus on the computer screen and the beginning of the emission of its pronunciation by the subject, were used test the hypothesis that the phonetic nature of the initial phonemes of the words affects the measurement of the RT, since for words beginning with stop consonants, for example, the sound itself contains an

average silence lasting 100ms. If the phonetic nature of the first sound influences the measurement of the RT, this means that in tasks of reading aloud, this phenomenon should be described and controlled. Thus, a list of 58 words varying in the phonetic nature of their first sound – vowel, stop (voiced and unvoiced), fricative (voiced and unvoiced), nasal and lateral – were presented to children from 1st to 4th grades and to adult readers, with their reaction times recorded. The first results do not confirm the hypothesis that the silence of the stop sound is affecting the RT, since no significant difference was found between the average RT between the stops and fricatives. The shortest RTs were recorded for the nasal consonants and the vowels (*o*, *a*) for the adult group and for all the grades (excepting the vowel *a* for the first grade). The longest RTs were recorded for the fricative sound R (except in the third grade), and *u* (except in the adult group). The behavior of the sound *u* was distinctly curious, not following the pattern of the other vowels in the children's reading.

As tarefas de leitura em voz alta de palavras isoladas pertencem à área de estudo da Psicologia Cognitiva conhecida como “Reconhecimento de palavras”. Consistem na apresentação de listas de palavras reais e de não-palavras, para a leitura. As listas de palavras reais variam em freqüência de ocorrência (palavras, de alta e baixa freqüência) em regularidade de correspondência letra-som (palavras regulares e irregulares) e em comprimento (número de sílabas). As listas de não-palavras, usualmente com igual número de estímulos das listas de palavras reais, são construídas com a mesma estrutura ortográfica e os mesmos comprimentos dos estímulos usados nas listas de palavras reais. Freqüentemente, as palavras reais e não-palavras são testadas em blocos separados. O tempo de reação e os erros cometidos são registrados.

A medida de tempo de reação é feita com base no tempo gasto entre a apresentação de um estímulo na tela de um computador e o início da emissão da resposta pelo sujeito. Nesses experimentos, o início da onda sonora produzida pelo sujeito é captado por um microfone e enviado para um cronômetro dentro do computador que capta o início da fala e mede o tempo de reação em milésimos de segundos. Um ponto importante, que tem passado despercebido pelos pesquisadores, é que a natureza fonética dos fones iniciais das palavras pode afetar a medida de tempo de reação. Diante desta possibilidade, nossa preocupação, de ordem metodológica, tem por objetivo investigar o efeito da natureza fonética do fone inicial da palavra no tempo de reação medido, quando se lê a palavra. A questão é, então, a seguinte: a natureza fonética dos fones iniciais das palavras afeta a medida de tempo de reação? O conhecimento do efeito do fone inicial das palavras na medida de tempo de reação

possibilitará aos pesquisadores aumentar o controle de variáveis externas que possam afetar a precisão da medida de tempo de reação nos experimentos de leitura em voz alta e desta forma aumentar a confiabilidade dos resultados obtidos.

Assim, por exemplo, como a fase de silêncio da oclusiva (por volta de 100 ms) não é identificada pelo sistema, as palavras iniciadas por sons oclusivos seriam lidas com um tempo de reação maior do que para os outros fones, já que estes não são precedidos por uma fase de silêncio. O que se espera, então, é que as palavras iniciadas por oclusivas apresentem tempos de reações anormalmente mais longos. Conforme podemos observar nas figuras abaixo, o oscilograma das palavras “tapa”, “fama” e “mala” foi sincronizado no seu ponto inicial, a partir de 100 ms no eixo do tempo, observando-se uma ausência de energia nos primeiros 125 ms da palavra tapa, o que não ocorre com as duas outras palavras.

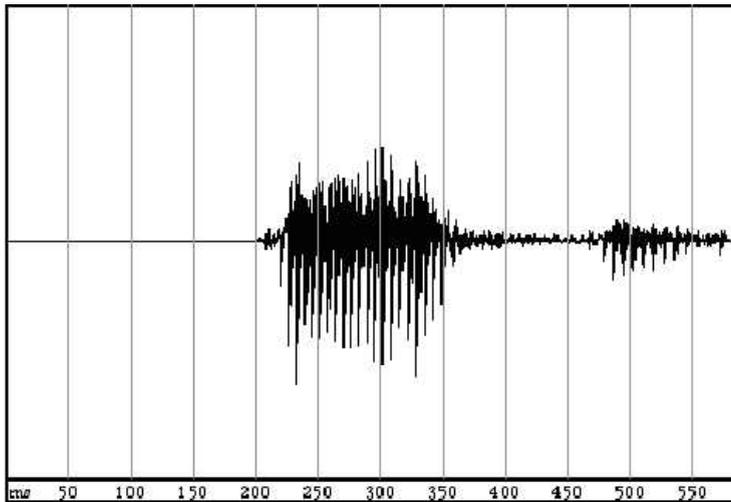


Figura 01: Oscilograma da palavra “tapa”.

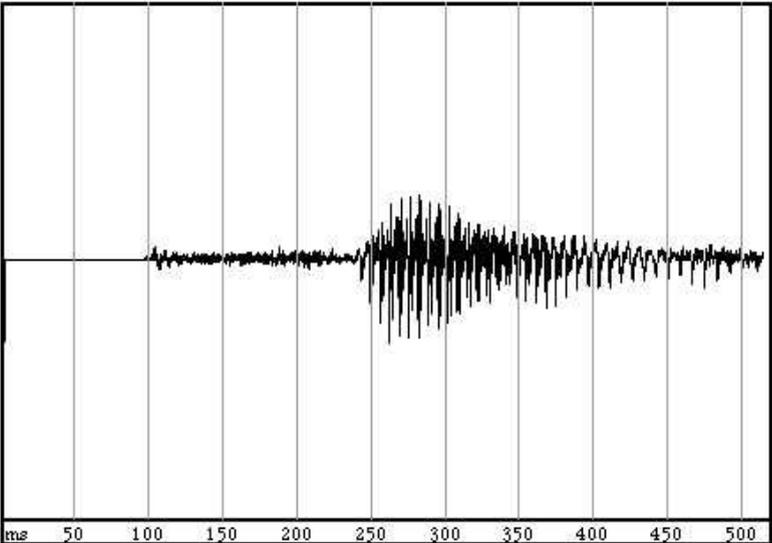


Figura 02 Oscilograma da palavra “fama”.

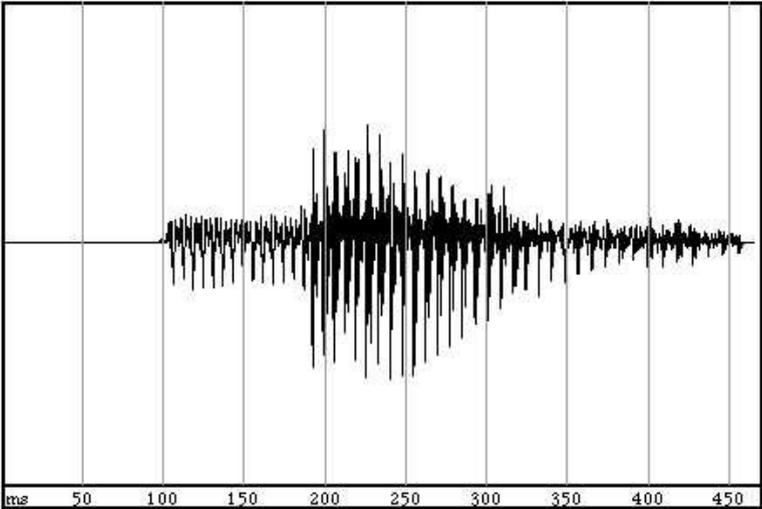


Figura 03: Oscilograma da palavra “mala”.

A confirmação da hipótese de que a fase de silêncio da oclusiva pode afetar a medida de tempo de reação terá um impacto direto na construção dos estímulos que irão compor as listas de palavras usadas nas tarefas de leitura em voz alta. Para controlar o maior tempo de reação na leitura de palavras iniciadas por sons oclusivos, as listas experimentais terão de ser construídas de forma que, além dos controles normalmente usados – nível de frequência de ocorrência, nível de regularidade grafema-fonema e estrutura silábica das palavras –, seja necessário a introdução de um controle adicional, nesse caso, a natureza do som inicial das palavras. Na prática, isso significaria que, nas tarefas de leitura em voz alta, as palavras de alta e baixa frequência (AF e BF), assim como as regulares (R) e irregulares (I) dentro de cada nível de frequência, deveriam começar com o mesmo som, como ilustrado abaixo:

Tabela 1: Palavras de alta frequência (AF) e baixa frequência (BF); regulares (R) e irregulares (I).

AF	R	NADA
	I	NOVO
BF	R	NAVE
	I	NORA

Assim, qualquer diferença no processamento de cada classe de palavras poderia, com maior segurança, ser atribuída a fatores inerentes à variável em questão e não a fatores estranhos. Alternativamente, o tempo de silêncio da oclusiva deverá ser computado ao tempo de reação das palavras iniciadas por oclusivas e levado em conta na interpretação dos resultados. No entanto, se for descoberto que a natureza fonética do primeiro fone da palavra não afeta a medida de tempo de reação, o pesquisador não precisará levar em conta mais essa variável na escolha das palavras a serem lidas.

MÉTODO

Amostra

A amostra constituiu-se de dois grupos de sujeitos. O primeiro foi formado por 80 crianças da 1ª à 4ª série do ensino fundamental (20 crianças por série) matriculadas no Centro Pedagógico da UFMG, em Belo Horizonte – todas apresentando desenvolvimento normal na habilidade de leitura, segundo o julgamento da professora. O segundo grupo, introduzido para fins de controle, foi composto por 29 alunos do 5º período do curso de Psicologia da UFMG.

Instrumento

O instrumento foi constituído de uma lista de 58 palavras, variando na natureza fonética de seu som inicial – vogal, oclusiva (vozeada e não vozeada), fricativa (vozeada, não vozeada), nasal, e lateral. A lista incluiu, ainda, um grupo com a fricativa mais posterior do português, que no nosso dialeto é mais comumente pronunciada glotal ([h] e [ɦ]), representada pela letra ‘R’, tendo em vista sua variabilidade articulatória (ver Tabela 2). Todas as palavras eram compostas de 4 letras, regulares para a leitura e de baixa frequência de ocorrência para as crianças da 1ª à 4ª série (de acordo com classificação de PINHEIRO, 1996, 2003). Optamos por palavras regulares e com estrutura silábica predominantemente CVCV como estímulo deste estudo para garantir uma leitura mais eficiente o que nos possibilitaria observar, com maior clareza, o impacto da variável independente – a natureza fonética do som inicial das palavras – na medida de TR. A escolha de palavras de baixa frequência de ocorrência se deveu ao fato de apenas nesse nível de frequência ser possível encontrar, em cada categoria de som desejada, um número suficiente de palavras regulares. Ainda assim a lista não apresenta palavras iniciadas com /i/ ou /e/ por esses estímulos, na categoria de palavra regular, não constarem em número suficiente no corpo de palavras de referência.

Tabela 2: Lista de palavras para o teste sobre a influência do fone inicial no TR para a leitura de palavras de 4 letras, regulares para a leitura e de Baixa Frequência de ocorrência para crianças da 1ª à 4ª série do ensino fundamental

Modo articulação e fonação*	OCL-NV	OCL-V	FRC-NV	FRC-V	NSL
Nº de itens por categoria	(7)	(7)	(7)	(7)	(7)
	PICO	BALA	FAMA	VAGA	MALA
	PURA	BIFE	FIGO	VALA	MICO
	TABA	BULE	FUGA	VARA	MINA
	TAPA	DAMA	FUMO	ZEBU	MUDO
	TUBA	DURO	SAPA	JACA	MULA
	CAVA	GAGO	SINO	JATO	NAVE
	CUCA	GAZE	SUJO	JIPE	NABO
Modo articulação e fonação	R	LRL	A	O	U
Nº de itens por categoria	(4)	(4)	(5)	(5)	(5)
	RAMO	LAJE	ALVO	OITO	UIVO
	RARA	LAMA	ANIL	ONDA	UNIR
	RISO	LISO	ANJO	ONZE	URNA
	RUMO	LUTA	ARMA	ORAL	UVAS
			ATOR	OURO	UMAS

OCL-NV – oclusiva não-vozeada

OCL-V – oclusiva vozeada

FRC-NV – fricativa não-vozeada

FRC-V – fricativa vozeada

NSL – nasal

LRL – lateral

A – vogal /a/

O – vogal /o/

R – fricativo glotal

U – vogal /u/

Procedimento

A lista de palavras foi administrada por meio de um microcomputador portátil, dirigido pelo programa *Mel Professional* que possui recursos para medir o tempo de reação (TR – tempo entre o aparecimento do estímulo e emissão da pronúncia da palavra), em milésimos de segundo e controlar intervalos entre estímulos. Os itens da lista foram apresentados em letras minúsculas, no centro da tela do monitor, e foram vistos a uma distância de 50 a 75 centímetros,

dependendo dos movimentos do sujeitos. A cada tentativa, um estímulo era apresentado no centro da tela do micro, precedido de um sinal de alerta. Assim que o sujeito pronunciava a palavra, o TR era marcado.

O estudo foi realizado nos meses maio-junho 2001 e o teste foi aplicado por bolsistas do projeto PAD e PIBIC no laboratório de Psicologia do DECAI da FAE/UFMG. Cada criança foi testada individualmente em uma sessão. Inicialmente, explicou-se à criança como executar a tarefa, o que foi demonstrado por um treinamento que foi introduzido por meio da seguinte instrução: “após um sinal de alerta – três estrelinhas – você vai ser apresentado a uma lista de palavras escritas que aparecerão no centro da tela, uma de cada vez. Tente ler cada palavra o mais rápido e corretamente que puder”. A lista de palavras do treinamento consistia de 15 itens. Após o treinamento, o teste propriamente dito foi introduzido. Para cada sujeito, os itens da lista experimental eram embaralhados pelo computador. As respostas foram gravadas e classificadas como correta, errada ou anulada (em caso de falha do equipamento) pelo aplicador.

RESULTADOS

Os resultados consistiram em medidas de TR para as respostas corretas para a leitura das 58 palavras que compunham os estímulos do estudo. Como esperado, os sujeitos de todos os grupos leram os estímulos com alto índice de precisão. Para cada nível acadêmico estudado, obteve-se a média de respostas corretas para cada classe de palavras. A seguir, as médias das variáveis contrastantes – por exemplo, palavras iniciadas por oclusivas vozeadas e oclusivas não-vozeadas – foram comparadas por t-testes. Todas as comparações efetuadas são apresentadas na Tabela 3.

Tabela 3: Comparações efetuadas em pares de médias de TR para as respostas corretas para cada grupo de sujeitos. As comparações de maior interesse encontram-se destacadas

OCLV x OCLNV	OCLV+OCLNV x FRCV+FRCNV	A x OCLV	O x OCLV	U x OCLV
FRCV x FRCNV	OCLV+OCLNV x NSL	A x OCLNV	O x OCLNV	U x OCLNV
	OCLV+OCLNV x R	A x FRCV	O x FRCV	U x FRCV
OCLV x FRCV	OCLV+OCLNV x LRL	A x FRCNV	O x FRCNV	U x FRCNV
OCLV x NSL		A x NSL	O x NSL	U x NSL
OCLV x R	FRCV+FRCNV x NSL	A x R	O x R	U x R
OCLV x LRL	FRCV+FRCNV x R	A x LRL	O x LRL	U x LRL
	FRCV+FRCNV x LRL			
OCLNV x FRCNV		A x O	O x U	
OCLNV x NSL	R x LRL	A x U		
OCLNV x R				
OCLNV x LRL				

A Tabela 4 (A-F) apresenta, para os grupos estudados, os resultados das comparações entre categorias de palavras que geraram diferenças significativas entre médias de tempo de reação para as respostas corretas.

Os resultados não confirmam a hipótese de o silêncio da **oclusiva** estar afetando o TR, já que em nenhum grupo as oclusivas foram lidas com TR significativamente maior do que as fricativas. As únicas diferenças significativas encontradas nas médias de TR entre as oclusivas e as fricativas ocorreram na 2ª e 4ª séries, na condição não-vozeada, mas, ao contrário do esperado, as oclusivas foram de fato lidas com menor TR do que as fricativas.

No que se refere à fonação das consoantes, diferenças significativas entre o TR de consoantes vozeadas e não vozeadas foram encontradas apenas entre as oclusivas, com as consoantes não vozeadas apresentando, a partir da segunda série, consistentemente menor TR do que as vozeadas. Na 2ª e 4ª séries e no 5º período, a diferença encontrada foi significativa. O efeito da fonação das consoantes na medida de TR parece, pois, restrito às oclusivas e tende a ser evidenciado com o aumento de proficiência na leitura.

Tabela 4 (A-F). Médias (M) de TR (em ms) e desvio padrão (DP) para as respostas corretas de pares de fonemas em início de palavras cuja diferença de TR foi significativa na leitura de crianças da 1ª à 4ª série e de alunos do 5º período do Curso de Graduação de Psicologia da UFMG.

A) crianças da 1ª série - N^o = 20

Par	M do 1º membro do par	DP	M do 2º membro do par	DP	Diferença	Test-t	Nível de signif.
U – A	1352	447	1219	371	133	(19)-2,299	0,033*
U – OCLNVN ¹	1352	447	1215	387	137	(19)-2,299	0,005**
U – OCLV	1352	447	1211	344	141	(19)-2,793	0,012*
U – LRL	1352	447	1191	459	161	(19)-2,224	0,038*
U – FRV	1352	447	1155	297	225	(19)-3,913	0,001***
U – FRNV	1352	447	1127	252	225	(19)-3,398	0,003**
U – O	1352	447	1087	281	265	(19)-4,814	0,000***
U – NSL	1352	447	1054	250	298	(19)-4,563	0,000***
R – FRVNV ²	1324	478	1140	261	184	(19)-2,219	0,039*
R – O	1324	478	1087	281	237	(19)-2,934	0,009**
A – O	1219	371	1087	281	132	(19) 2,509	0,021*
A – NSL	1219	371	1054	250	165	(19) 3,241	0,004**
OCLVNV – NSL	1215	387	1054	250	161	(19)-3,352	0,003**
OCLV – O	1211	344	1087	281	124	(19)-2,937	0,008**
OCLV – NSL	1211	344	1054	250	157	(19)-4,331	0,000***
FRV – O	1155	297	1087	281	68	(19) 2,698	0,014*
FRVNV – O	1140	261	1087	281	56	(19) 2,643	0,016*
FRVNV – NSL	1140	261	1054	250	86	(19) 3,631	0,002**

* = > 0,05; ** = > 0,01; *** = > 0,001

¹ Consoantes oclusivas vozeadas e não-vozeadas juntas.

² Consoantes fricativas vozeadas e não-vozeadas juntas.

B) crianças da 2ª série - N^o = 20

Par	M do 1º membro do par	DP	M do 2º membro do par	DP	Diferença	Test-t	Nível de signif.
U – FRVNV	1171	356	1047	322	124	(19)-2,350	0,030*
U – OCLV	1171	356	1020	266	151	(19)-3,002	0,008**
U – A	1171	348	975	361	196	(19)-2,475	0,023*
U – OCLNV	1171	348	963	293	208	(19)-3,638	0,002**
U – LRL	1171	356	960	253	211	(19)-4,215	0,001***
U – O	1171	348	948	320	223	(19)-4,103	0,001***
U – NSL	1171	356	908	230	263	(19)-5,445	0,000***
R – OCLV	1124	386	1020	266	104	(19)-2,065	0,54*
R – OCLVNV	1124	386	991	278	133	(19)-2,612	0,018*
R – A	1124	386	975	345	149	(19)-3,326	0,004**
R – OCLNV	1124	386	963	301	161	(19)-3,067	0,007**
R – LRL	1124	386	960	254	164	(19)-3,088	0,006**
R – O	1124	386	948	328	176	(19)-3,378	0,003**
FRV – O	1077	391	948	328	129	(19) 2,716	0,014*
FRVNV – NSL	1047	322	908	230	139	(19) 3,861	0,001***
OCLV – OCLNV	1021	266	963	301	58	(19)-2,835	0,011*
OCLV – NSL	1021	266	908	230	113	(19)-4,339	0,000***
OCLV – O	1021	266	948	328	78	(19)-2,333	0,038*
FRNV – OCLNV	1018	270	963	301	55	(19)-2,248	0,037*

C) crianças da 3ª série - N^o = 20

Par	M do 1º membro do par	DP	M do 2º membro do par	DP	Diferença	Test-t	Nível de signif.
U – OCLV	1112	500	1016	451	96	(19)-2,399	0,027*
U – FRV	1112	500	1015	512	97	(19)-2,986	0,008**
U – FRVNV	1112	500	992	424	120	(19)-3,458	0,003**
U – R	1112	500	985	297	127	(19)-2,051	0,054*
U – OCLNV	1112	500	977	542	135	(19)-3,187	0,005**
U – NSL	1112	500	939	424	137	(19)-3,786	0,001***
U – A	1112	500	935	425	177	(19)-3,345	0,004**
U – A	1112	508	931		181		0,000***
OCLV – NSL	1016	451	939	424	77	(19)-2,255	0,036*
OCLV – A	1016	455	935	425	81	(19)-2,861	0,015*
OCLV – O	1016	451	931	472	85	(19)-2,899	0,009**
FRV – O	1015	512	931	472	84	(19)-2,357	0,029*
OCLVNV – NSL	999	485	939	424	60	(19)-2,364	0,029*

D) crianças da 4ª série - N^o = 20

Par	M do 1º membro do par	DP	M do 2º membro do par	DP	Diferença	Test-t	Nível de signif.
R – OCLV	1057	427	856	148	201	(19)-2,840	0,010**
R – FRVNV	1057	427	851	154	206	(19)-2,991	0,008**
R – OCLVNV	1057	427	827	146	230	(19)-3,264	0,004**
R – LRL	1057	427	795	185	262	(19)-3,068	0,006**
R – OCLNV	1057	427	793	158	264	(19)-3,671	0,002**
R – A	1057	427	778	125	279	(19)-3,292	0,004**
R – O	1057	427	747	130	310	(19)-3,403	0,003**
U – FRV	978	233	859	185	119	(19)-2,396	0,027*
U – OCLV	978	233	856	148	122	(19)-2,467	0,023*
U – FRVNV	978	233	851	154	127	(19)-2,803	0,011*
U – OCLVNV	978	233	827	146	151	(19)-3,399	0,003**
U – LRL	978	233	795	185	183	(19)-4,631	0,000***
U – NSL	978	233	788	146	190	(19)-4,774	0,000***
U – A	978	233	778	125	200	(19)-5,136	0,000***
U – O	978	233	747	130	231	(19)-5,710	0,000***
FRV – A	859	185	778	125	81	(19)-2,923	0,009**
FRV – O	859	185	747	130	112	(19)-3,176	0,005**
OCLV – OCNV	856	148	793	158	63	(19)-3,068	0,006**
OCLV – NSL	856	148	788	146	68	(19)-3,223	0,004**
OCLV – A	856	148	778	125	78	(19)-2,893	0,009**
OCLV – O	856	148	747	130	109	(19)-3,945	0,001***
FRVNV – NSL	851	153	788	146	63	(19)-3,075	0,006**
FRNV – OCNV	844	146	793	158	51	(19)-2,576	0,018*
FRNV – O	844	146	747	130	97	(19)-3,730	0,001***
FRNV – A	844	146	778	125	66	(19)-2,458	0,024*
OCLVNV – NSL	827	146	788	146	39	(19)-2,490	0,022*

E) Alunos do 5º período do Curso de Graduação de Psicologia - UFMG
Nº = 29

Par	M do 1º membro do par	DP	M do 2º membro do par	DP	Diferença	Test-t	Nível de signif.
FRV – A	533	82	498	78	34	(28) -20,06	0,000***
FRV – O	533	82	488	81	45	(28) -4,852	0,000***
FRNV – A	532	88	498	78	34	(28) -3,674	0,001***
FRNV – O	532	88	488	81	44	(28) 5,422	0,000***
FRNVN – NSL	532	81	511	93	21	(28) -3,478	0,002**
OCLV – OCNV	532	77	516	70	16	(28) -2,343	0,026*
OCLV – NSL	532	77	511	93	21	(28) -2,533	0,017*
OCLV – A	532	77	498	78	34	(28) -3,854	0,001***
OCLV – O	532	77	488	81	44	(28) -6,223	0,000***
R – O	520	87	488	81	32	(28) -3,248	0,003**
OCLNV – A	516	70	498	78	18	(28) -2,847	0,008**
OCLNV – O	516	70	488	81	28	(28) -4,692	0,000***
U – O	515	91	488	81	27	(28) -3,118	0,004**
NSL – O	512	93	488	81	24	(28) -2,901	0,007**
LRL – O	512	107	488	81	24	(28) -2,147	0,041*

A Tabela 5 apresenta, para os dois grupos de sujeitos, a ordem decrescente de médias de TR para cada categoria de palavras.

Tabela 5 - Ordem decrescente das medias de TR para cada categoria de estímulos para crianças da 1ª à 4ª série e alunos do 5º período do Curso de Psicologia/UFMG

1ª Série		2ª Série		3ª Série		4ª Série	
Fone	TR	Fone	TR	Fone	TR	Fone	TR
U	1352	U	1171	U	1112	R	1057
R	1324	R	1124	OCV	1016	U	997
OCNV	1220	FRV	1077	FRV	1015	FRV	859
A	1219	FRVNV	1047	LRL	1004	OCV	856
OCVNV	1215	OCV	1021	OCVNV	999	FRVNV	851
OCV	1211	FRNV	1018	FRVNV	992	FRNV	844
LRL	1191	OCVNV	990	R	985	OCVNV	827
FRV	1155	A	975	FRNV	979	LRL	795
FRVNV	1140	OCNV	963	OCNV	977	OCNV	793
FRNV	1127	LRL	960	NSL	939	NSL	788
O	1087	O	948	A	935	A	778
NSL	1054	NSL	908	O	931	O	747

5º período

Som	TR
FRV	533
FRNV	532
FRVNV	532
OCV	532
OCVNV	524
R	520
OCNV	516
U	515
NSL	512
LRL	512
A	498
O	488

Exceto na 1ª série para a vogal **a**, as **consoantes nasais** e as **vogais (o, a)** apresentaram o menor TR em todas as séries. Os fones **R** (fricativo glotal) e **u** apresentaram maior TR (exceto no grupo de adultos). A **lateral**, exceto na 3ª série, tendeu a apresentar TRs iguais ou levemente acima dos TRs apresentados pelas nasais e vogais. O comportamento do **u** – no grupo de crianças – mostrou-se bastante curioso, pois não acompanhou as demais vogais. Uma explicação seria a sua baixa ocorrência no início das palavras expostas a crianças. Uma evidência dessa possibilidade é que na leitura dos adultos o TR para as palavras iniciadas com o **u** não diferiu significativamente do TR das nasais, laterais e do **a**. Isso sugeriu, portanto, que no leitor inteiramente hábil, esta vogal é processada de forma mais rápida do que na fase de aquisição e desenvolvimento da leitura.

Um ponto mais geral, mas que merece menção por confirmar resultados anteriores sobre o desenvolvimento da habilidade de reconhecimento de palavras (por exemplo, PINHEIRO, 1994, 1995), é o declínio constante do tempo de reação observado na leitura de todos os estímulos com o avanço da escolarização.

CONCLUSÕES

A pouca preocupação com o controle da influência da natureza fonética do fone inicial dos estímulos usados nos experimentos de leitura em voz alta na medida de TR, observada na literatura, parece injustificada. Embora a hipótese principal do estudo não tenha se confirmado, os resultados obtidos mostram que, pelo menos nos estudos com crianças, a escolha das palavras a serem usadas nos experimentos de leitura em voz alta requer alguns cuidados no que se refere ao fone inicial dos estímulos usados (além dos controles tradicionais). Por exemplo, uma lista composta de muitas palavras de alta frequência começadas com a vogal **u** e com a consoante **R** (sons que apresentaram, na leitura das crianças, tempos de reações mais altos) e de muitas palavras de baixa frequência começadas com as vogais **a**, **o** e com **nasais** (sons que apresentaram menores tempos de reações na leitura da mesma população) pode, no mínimo, diminuir

a magnitude do efeito de frequência (leitura mais rápida e precisa de palavras de alta frequência em relação a palavras de baixa frequência), normalmente encontrado nas tarefas de leitura em voz alta, comprometendo, assim, os resultados obtidos. Além disso, um TR maior para as oclusivas em relação às fricativas ocorreu apenas para a primeira série, enquanto no grupo controle, de universitários, o TR para as consoantes praticamente não variou. Tomando-se apenas esses resultados, pode-se, talvez, concluir que a natureza fonética do fone inicial devesse ser necessariamente controlada nos testes para a primeira série, quando, certamente, o uso da rota fonológica é mais comum.

Parece, pois, que conseguimos identificar um ponto teórico no estudo de reconhecimento de palavras o qual merece um pouco mais de atenção dos pesquisadores da área. O estudo aqui relatado representou apenas um primeiro passo para o entendimento do fenômeno sob análise. Experimento com controle de outras variáveis, como a estrutura silábica, o grau de frequência da palavra, assim como as características do equipamento de registro de TR poderá esclarecer melhor se a natureza fonética do fone inicial da palavra afeta a TR da palavra.

REFERÊNCIAS BIBLIOGRÁFICAS

PINHEIRO, A. M. V. Reading and spelling development in Brazilian Portuguese. *Reading & Writing*, special issue on Literacy Acquisition, 7/1 mar., p. 111-138, 1995.

PINHEIRO, A. M. V. *Contagem de frequência de ocorrência e análise psicolinguísticas de palavras expostas a crianças na faixa pré-escolar e séries iniciais do 1º grau*. São Paulo: Associação Brasileira de Dislexia (ABD), 1996.

PINHEIRO, A. M. V. *Vocabulário comum de alta, média e baixa frequência de ocorrência para crianças da 1ª à 4ª série do ensino fundamental*. Relatório Final Global e Integrado de atividades desenvolvidas no projeto "Avaliação cognitiva das capacidades de leitura e de escrita de crianças nas séries iniciais do 1º grau – AVACLE", aprovado pelo CNPq em 2003.

PINHEIRO, A. M. V. *A leitura e a escrita: uma abordagem cognitiva*. 2. ed. Campinas: Editora Livro Pleno, 2004.

Analysis of acoustic-prosodic features of spontaneous expressive speech

Carlos Toshinori Ishi
Nick Campbell

JST/CREST ESP Project at ATR/HIS Labs. - Japan

Resumo

A análise de características acústico-prosódicas foi conduzida em um banco de dados de fala espontânea, com o propósito de construir um sintetizador de fala capaz de sintetizar uma fala expressiva, ou seja, vários estilos de fala. Os estilos de fala foram categorizados de acordo com os seguintes itens: estado emocional percebido (neutro, preocupado, contente, insatisfeito, bravo, cansado, etc.); atitude com respeito ao ouvinte (grau de polidez, gentileza, consideração); atitude com respeito ao conteúdo da fala (grau de confiança, envolvimento, interesse); qualidade da voz (energia, brilho, dureza). As seguintes características acústicas foram analisadas: F0

e RMS (relacionados à altura e à intensidade da voz), terceiro e quarto formantes (F3 e F4, relacionados com o comprimento do trato vocal, e podem estar correlacionados com o grau de extensão dos lábios), e quociente de amplitude (AQ – “amplitude quotient”, relacionado com o tipo de fonação da voz). Análise de componentes principais e coeficientes de correlação foram usados para avaliar os dados. Os resultados indicaram correlação entre algumas categorias do estilo da fala e da qualidade de voz, por exemplo, energia e envolvimento, brilho e interesse, gentileza e dureza. Correlações significativas foram obtidas entre qualidade de voz e características acústico-prosódicas. Valores médios de F0, RMS, AQ e F4 foram encontrados como sendo os parâmetros mais influentes na qualidade de voz, indicando que uma combinação destas características acústicas pode, possivelmente, ser usada para prever qualidades de voz.

1. INTRODUCTION

For unrestricted text-to-speech conversion, the problems of text anomaly resolution and focus determination are profound. They can require a level of world-knowledge and discourse modeling that is still beyond the capability of most text-to-speech synthesis systems. One implication of this is that the prosody component of the speech synthesizer can only be provided with a default specification of the intentions of the speaker or the underlying discourse-related meanings of the utterance, resulting in a flat rendering of the text into speech. This is not a problem for the majority of synthesis applications, such as news-reading or information announcement services, but if the synthesizer is to be used in place of a human voice for interactive spoken dialogue, then the speech will be perceived as lacking illocutionary force, or worse, it will give the listener a false impression of the intention of the utterance, leading to potential misunderstandings.

When a synthesizer is to be used in place of a human voice in conversational situations, such as in a communication aid for the vocally impaired, or in call-centre operations, then there is a clear need for the vocal expression of more than just the semantic and syntactic linguistic content of the utterance.

Since the information carried by human speech includes linguistic, paralinguistic, and extra-linguistic details, the listener presumably parses all three sources to gain access to the intended meaning of each utterance. Paralinguistic information, signaled by tone-of-voice, and speaking style, becomes more important as the conversation becomes more personal. When speaking with friends, for example, we normally use a different speaking style and tone-of-voice than when addressing a stranger or a wider audience.

Speech synthesis must likewise become capable of expressing such differences.

The aim of this work is to investigate how speakers control prosodic features (like intonation, rhythm and voice quality), which are embedded in speech utterances, to express several levels of emotions, intentions and attitudes, and apply this knowledge for speech synthesis applications. For this purpose, it is necessary to investigate the relationship between the human impressions and the acoustic-prosodic features extracted from the speech signal.

2. SPEECH DATA AND SPEECH UNIT FOR ANALYSIS

There has been considerable attention paid recently to the construction and evaluation of language resources [1], because without well-constructed text and speech databases it will be difficult to make technological advances in the areas of speech and language processing. However, current speech databases are limited in their ability to illustrate supra-linguistic speech variation because of the constraints on corpus design.

There are several works concerning emotional speech using controlled “emotional” speech utterances for analysis, but there are doubts about the naturalness of the speech. More recently, the speech content has become less controlled. The latest corpus initiative in Japan [2] is focusing on prepared spontaneous speech produced in a monologue context, especially including conference and workshop presentations.

In the present work, we used part of the speech database recorded in the JST (Japan Science & Technology Agency) CREST (Core Research for Evolutional Science and Technology) ESP (Expressive Speech Processing) Project [3,4], as speech data for analysis. As for the speech data recording, volunteers recorded their daily spoken interactions to minidisc devices by wearing head-mounted close-talking microphones. These samples were then transcribed manually and segmentally aligned automatically from the transcriptions, using HMM-based techniques.

In the current paper, we used part of the speech database, containing everyday conversations of one Japanese adult female speaker. This speech data contains very natural conversations with family and friends, phone calls to company, hotel, and hospital.

As speech units for analysis, we decided to use prosodic phrases [5], because changes in speaking styles are common from phrase to phrase within the same sentence. Even though intra-phrase changes in speaking styles can also occur, it is less common than inter-phrase changes. Prosodic phrase boundaries were placed semi-automatically at long pauses and pitch resets.

3. CATEGORIZATION AND LABELING OF SPEAKING STYLES AND VOICE QUALITY

The speaking styles were categorized in terms of the following items:

- Perceived mood or emotional state:
Neutral, worried, content, happy, discontent, angry, sad, tired, depressed, etc.: subjects were freely allowed to choose or propose a proper adjective that better represents the emotional state perceived from the speech utterance;
- Speaker's attitude with respect to the listener:
Degree of politeness: sloppy, slightly sloppy, casual, slightly polite, polite, formal.
Degree of warmth, and degree of considerateness: scale from -2 to +2;
Degree of softness: scale from 0 to +2;
- Speaker's attitude with respect to the contents:
Degree of confidence, involvement, and interest: scale from -2 to +2;
Here, -2/+2 indicates "noticeable", -1/+1 indicates "slightly noticeable" and "0" indicates "non-noticeable".

As for voice quality, the following categories were labeled:

- Energy (global intensity of the speech): 5 step scale from weak voice to strong voice.
- Brightness (closely related to mood): 5 step scale from dark voice (bad mood) to bright voice (good mood).
- Tension (related to the vocal effort): 5 step scale from soft voice to tense voice.

The speaker's intentions were also labeled, but their categories are omitted in the present paper, pending detailed acoustic-prosodic analyses.

Three native speakers of Japanese labeled the above categories for 670 phrases. Subjects were asked to label the perceived features of the speech regardless of the actual emotion of the speaker.

4. ACOUSTIC-PROSODIC FEATURES

Japanese is a pitch accent language, so that the F0 movements of an utterance are largely influenced by the accent types of the words that comprise the utterance. For this reason, in the current paper, we focused our analysis on overall acoustic-prosodic features of the speech.

The following acoustic features were analyzed: F0, RMS, third and fourth formants (F3 and F4), and glottal amplitude quotient (AQ):

- Average of F0 values (f0avg), related to the overall pitch height of the utterance, and F0 range, related to the overall pitch movement along the utterance; although F0 range is usually calculated as the difference between the maximum and minimum value, we decided to use the difference between the maximum and the mean value (f0dif), since the estimation of the minimum value is not so robust.
- Average of power values (rmsavg), related to the overall intensity of the utterance, and power range (rmsdif) as in the homologous F0 parameter definition. Power values were calculated taking the logarithm of the RMS (root mean square) values of the speech.

- Average of F3 and F4 (f_{3avg} , f_{4avg}). These higher formants are more directly related to the vocal tract length than are the first two formants, and could also be associated with lip spreading which is frequent in bright voices. The formants were extracted using the algorithm of cepstrum-to-formant mapping proposed in [7]. F1 and F2 were not used in the analysis because they are more closely related with phone identity than with prosodic information.
- Average of AQ values (aq_{avg}), related to the breathiness of the voice. The AQ parameter is defined as the ratio of the peak-to-peak value of the glottal waveform (after eliminating the formant effects) and the maximum negative peak of the derivative of the glottal waveform; it has been shown to be related to perceived phonation quality along the pressed/breathy continuum [6], and may also be related to the perception of softness (tension) of the speech. In this study, the negative of AQ is used in order to obtain higher values for pressed phonation (or tense voice) and lower values for breathy phonation (or soft voice).

The values of the parameters in only the vowel portions were used in the computation of parameters defined above.

5. ANALYSIS RESULTS

Principal component analysis techniques and correlation coefficients were used to evaluate the data.

5.1. Analysis results for items related to speaking style

Figure 1 shows the distributions of the labeled emotional states. Neutral utterances were found to be predominant because most of the data are phone call conversations. Further, although there is a brightness category for voice quality, labelers preferred to classify “Neutral” emotional states in “Neutral(dark)” and “Neutral(bright)”.

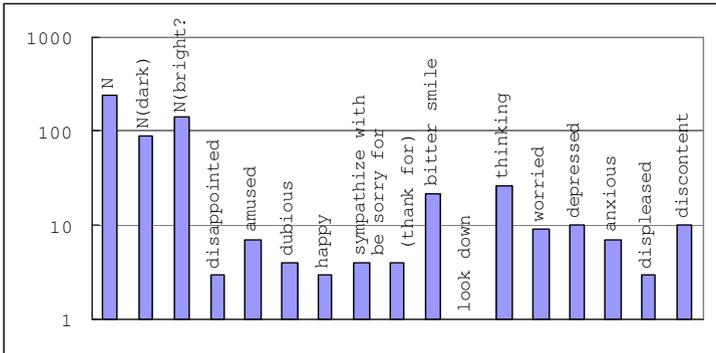


Figure 1. Distribution (number of occurrences) of the emotional states

Fig. 2 shows the distribution of the emotional states according to the politeness. From these distributions, it can be observed that {sloppy, casual} tend to be associated with {N(dark), depressed, discontent, displeased, bitter smile}, and that {casual, slightly polite} tend to be associated with {worried, anxious}. Furthermore, only {neutral, neutral(bright)} appear to be associated with the categories {polite, formal}.

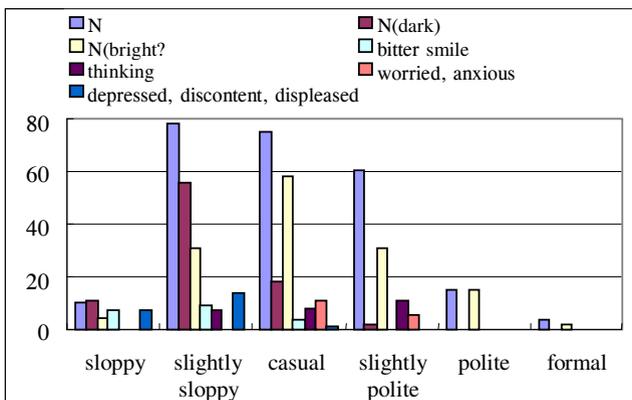


Figure 2 – Distribution (number of occurrences) of the emotional states according to politeness

Also, a strong relation was observed between the degree of politeness and the relationship between the speaker and the listener. For example, {casual, non-polite} manners were predominantly used for relatives and friends, while {polite, formal} manners were predominantly used for company, hospital and acquaintances.

Next, principal component analysis was conducted on the quantifiable categories related to speaking style. Fig. 3 shows the results in terms of the first two principal components. Table 1 shows the correlation coefficients between the categories.

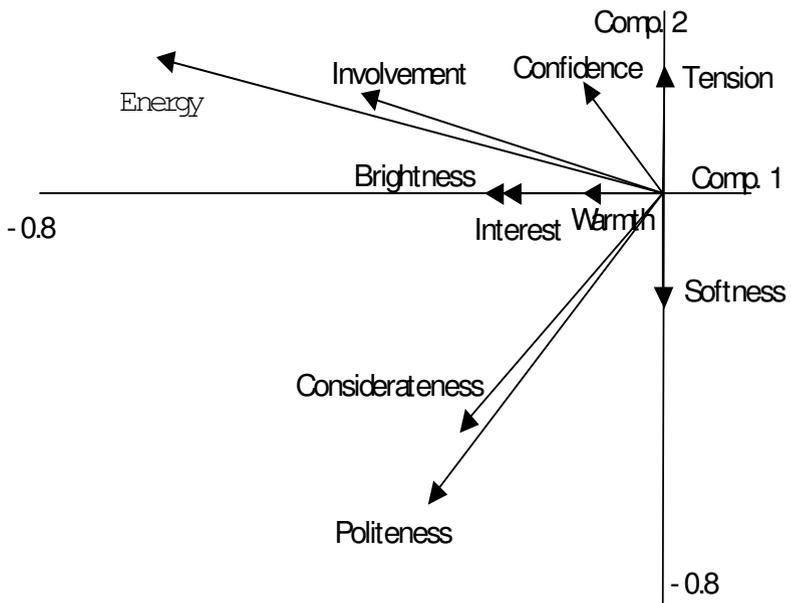


Figure 3 – Results of Principal Component Analysis on categories related to speaking styles

	Pol.	Warm	Soft	Cons.	Int.	Inv.	Conf.	Ener.	Tens.
Brightness	0.18	0.28	-0.02	0.18	0.29	0.39	0.26	0.46	0.12
Polite		0.13	0.28	0.63	0.20	0.20	-0.07	0.28	-0.17
Warmth			0.01	0.14	0.24	0.11	-0.01	0.25	-0.21
Soft (manner)				0.24	0.01	-0.05	-0.22	-0.12	-0.28
Considerate					0.33	0.21	-0.19	0.29	-0.12
Interest						0.42	0.17	0.50	-0.02
Involvement							0.41	0.72	0.31
Confidence								0.34	0.26
Energy									0.29

Table 1 – Correlation coefficient between the categories related to speaking style

The relationship between the categories related to speaking styles can be viewed in Fig. 3 and Table 1. Strong correlations were found mainly between “Energy” and “Involvement”, and between “Politeness” and “Considerateness”. Further, Figure 3 indicates a negative correlation between “Softness (manner)” and “Tension”; and a positive (but weak) correlation between “Brightness” and “Interest”. The small correlation coefficients found among most of the other categories indicate that they are relatively independent and can therefore be regarded as significantly separate perceptual labels.

5.2. Analysis of the acoustic-prosodic features

	<i>f0dif</i>	<i>rmsavg</i>	<i>rmsdif</i>	<i>f3avg</i>	<i>f4avg</i>	<i>aqavg</i>
<i>f0avg</i>	0.03	0.56	0.11	0.06	0.14	0.51
<i>f0dif</i>		0.26	0.32	-0.23	0.01	0.13
<i>rmsavg</i>			0.17	-0.24	0.06	0.51
<i>rmsdif</i>				-0.26	0.03	0.14
<i>f3avg</i>					0.12	-0.23
<i>f4avg</i>						0.04

Table 2 – Correlation coefficient between the acoustic-prosodic features

The results in Table 2 show correlations between $f0_{avg}$ and rms_{avg} , indicating that a high-pitched voice is usually uttered strongly. High correlations were also obtained between aq_{avg} and both rms_{avg} and $f0_{avg}$, indicating possibly that tense voice appears in high-pitched and high-energy voices. Further, the small correlation for the other features indicates that each feature carries different information.

5.3. Relation between the acoustic-prosodic features and the categories related to speaking styles

Fig. 4 shows the relationship among the acoustic-prosodic features and the voice quality categories in terms of the first and second principal components.

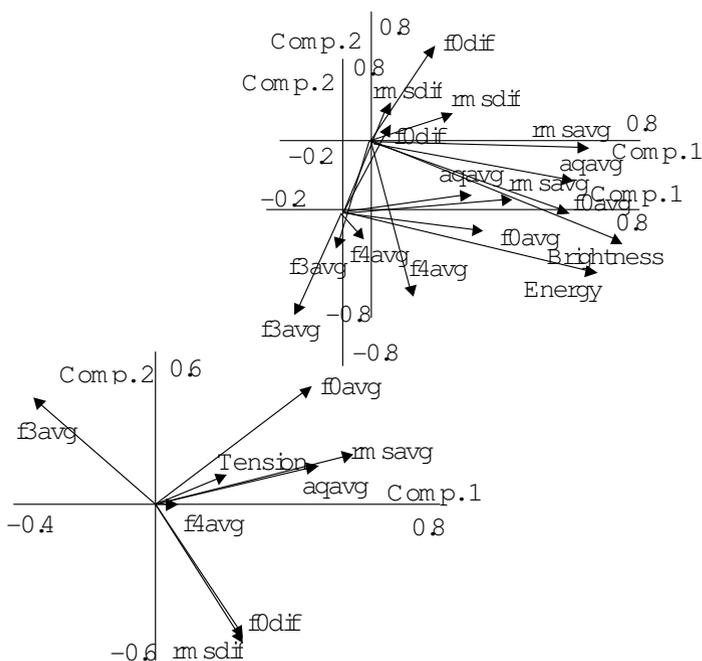


Figure 4 – PCA results for voice quality (“Energy”, “Brightness” and “Tension”) and acoustic-prosodic features

According to Fig. 4, “Energy” is correlated more with $f0_{avg}$ than with rms_{avg} . Similar tendencies can be seen for “Brightness”. However, a stronger component of $f4_{avg}$ was found for “Brightness” than for “Energy”. This result can be related to the tendency of good-mood speech to be uttered with spread lips, effectively shortening the vocal tract and thereby increasing $F4$. This correlation was not observed for $f3_{avg}$, indicating possibly that $F3$ is more affected by the vowel identity. As for the “Tension”, no strong correlation was found, but small correlations were found for rms_{avg} and aq_{avg} . Although AQ parameters are related to pressed or breathy phonation, other features may influence in the perception of “Tension”.

6. CONCLUSION

In the present paper, spontaneous expressive speech data were categorized according to the listener’s impressions for speaking styles and voice quality, and their relation with acoustic-prosodic features were analyzed.

Results showed correlations between some categories of speaking manners and voice quality, e.g., energy and involvement, brightness and interest, politeness and considerateness. Also, significant correlations were obtained between voice quality and acoustic-prosodic features. Averages of $F0$, RMS, AQ and $F4$ were found to be the most influential parameters on voice quality, indicating that a combination of these acoustic features can possibly be used to predict voice qualities. As a next step, we are investigating the dynamics of the acoustic-prosodic parameters.

ACKNOWLEDGEMENTS

We thank the JST institution for supporting this project. Also, we give special thanks for Dr. Parham Mokhtari for providing the AQ and formant parameters of all database.

REFERENCES

- [1] European Language Resources Association: www.elda.fr/catalog.html.
- [2] MAEKAWA, K.; KOISO, H.; FURUI, S.; ISAHARA, H. Spontaneous speech corpus of Japanese. *Proc. of LREC2000*, p. 947-952, 2000.
- [3] The JST/CREST Expressive Speech Processing Project, introductory web pages at: HYPERLINK "<http://www.isd.atr.co.jp/esp>", www.isd.atr.co.jp/esp.
- [4] CAMPBELL, N. The recording of emotional speech; JST/CREST database research. *Proc. of LREC2002*, v. 6, p. 2029-2032, 2002.
- [5] HIROSE, K.; FUJISAKI, H.; KAWAI, H. Generation of prosodic rule-synthesis of connected speech of Japanese. *Proc. of IEEE-ICASSP86*, p. 2415-2418, 1986.
- [6] MOKHTARI, P.; CAMPBELL, N. Perceptual validation of a voice quality parameter AQ automatically measured in acoustic islands of reliability. *Proc. of Acoustical Society of Japan Spring 2002*, v. I, p. 401-402, 2002.
- [7] BROAD, D. J.; CLERMONT, F. Formant estimation by linear transformation of the LPC cepstrum. *Journal of the Acoustical Society of America*, 86, v. 5, p. 2013-2017, 1989.

O uso de estímulos sintéticos em testes de percepção de vogais de uma língua estrangeira¹

Ricardo Augusto Hoffmann Bion
Universidade Federal de Santa Catarina

Abstract

The development of models like Flege's Speech Learning Model and Best's Perceptual Assimilation Model directed Interphonology research to the perception of target language contrasts. These models have triggered the development of perception tests, often using synthesized stimuli. The present paper discusses the ways previous studies have used these stimuli in vowel perception tests, and also presents the results of a pilot test developed in Brazil. While previous tests were concerned with checking the acoustic cues used in the perception of vowel contrasts, the present paper analyzes a test that can access the categorical boundaries in the perception of English front vowels by Brazilian Portuguese speakers and native English speakers. The stimuli for this test were created with the ASL synthesizer; innovations can be observed in the test development and application.

1. INTRODUÇÃO

Após o desenvolvimento de modelos como o *Speech Learning Model (SLM – Modelo de Aprendizagem da Fala)* de Flege (1995) e o *Perceptual Assimilation Model (PAM – Modelo de Assimilação Perceptiva)* de Best (1995), pôde-se observar uma importante mudança metodológica nas pesquisas na área da interfonologia (o sistema de regras fonológicas da língua estrangeira (LE) em desenvolvimento). Essas pesquisas, antes direcionadas para o estudo da produção dos sons da língua alvo, voltam-se agora para o estudo da percepção, tendo em vista que as hipóteses tanto do SLM quanto do PAM baseiam-se nessa habilidade.

Ambos modelos propõem que a eficácia em *perceber* contrastes não nativos está parcialmente determinada pela maneira através da qual estes fones são assimilados em relação às categorias fonéticas nativas. No PAM, sugere-se que os sons da LE são percebidos de acordo com suas semelhanças em relação aos sons da língua materna (LM) mais próximos articulatoriamente. O SLM difere-se do PAM, pois trata diretamente da aprendizagem da LE. Nele, argumenta-se que um mecanismo chamado *classificação por equivalência* dificultaria a criação de novas categorias fonéticas para sons da LE semelhantes aos da LM, mas assim não o faria com os sons bem distintos.

Ao tomar as hipóteses desses modelos como referência, torna-se necessário o desenvolvimento de testes que acessem esta habilidade (percepção) na língua alvo. Os estímulos naturalmente produzidos seriam os mais indicados para este propósito, não fosse sua aplicação prática tão limitada. Seu uso, em testes como o *Categorical Discrimination Test* (FLEGE, 1996?) por exemplo, pode indicar se se adquiriu ou não um contraste, mas não ajuda quando se objetiva verificar os limites categóricos em sua percepção ou apontar quais parâmetros acústicos estão sendo utilizados pelo aprendiz nessa discriminação.

Para esses objetivos de pesquisa, é indispensável a utilização de estímulos cujos parâmetros foram manipulados através de programas de síntese. Esses estímulos sintéticos vêm sendo amplamente utilizados nas pesquisas voltadas para o estudo da percepção das consoantes, onde muito se tem trabalhado com mudanças no *voice onset time* (o tempo entre a soltura da oclusão e o início do vozeamento da vogal seguinte) das consoantes oclusivas (FLEGE & EFTING, 1988; CARAMAZZA, A. et al., 1973). Todavia, nos testes de percepção de vogais, sua utilização ainda não foi suficientemente explorada, havendo uma grande necessidade de estudos que analisem suas diferentes possibilidades de uso.

Tendo isso em vista, pretende-se, no presente trabalho, discorrer sobre a precedência do uso desses estímulos sintéticos em testes de percepção de vogais de uma língua estrangeira, divulgar os resultados e avaliar as potencialidades de um teste de percepção piloto desenvolvido no Brasil.

2. ESTUDOS PRECEDENTES

Embora perca-se um pouco da naturalidade com o uso de estímulos sinteticamente produzidos/modificados, há precedência para seu uso em pesquisas sobre a percepção das vogais de LE. A maioria dessas pesquisas têm como objetivo comparar os parâmetros acústicos utilizados para a discriminação das vogais entre falantes nativos e não nativos de uma língua. A inclusão de falantes nativos tem a vantagem de demonstrar que, apesar de sintéticas, as vogais dos testes podem ser percebidas naturalmente.

Escudero (2000, 2002) investigou a discriminação das vogais inglesas /i/ e /i/ por falantes de espanhol com sete contínuos de vogais sintetizadas, nos quais eram variadas duração e/ou qualidade. A conclusão foi que o aprendiz de uma LE pode aprender a perceber um novo contraste que não exista em sua LM, embora os parâmetros de identificação possam diferir dos utilizados pelos falantes nativos: a duração foi a pista mais utilizada na discriminação pelos falantes

do espanhol, enquanto os falantes nativos dependeram mais da qualidade espectral.

Morrison (2002a, 2002b) comparou, em um estudo longitudinal, a percepção do mesmo contraste entre falantes nativos de inglês e aprendizes japoneses e espanhóis. Foi utilizado, nesse estudo, um contínuo multidimensional de fala sintetizada das palavras *beat* /bit/, *bit* /bit/, *bead* /bid/ e *bid* /bid/ variando em duração e qualidade. Os sujeitos foram testados depois de um e seis meses de sua chegada nos Estados Unidos. Enquanto os japoneses apresentaram um limite categórico baseado em informação espectral nos testes inicial e final, os falantes de espanhol não apresentaram limites categóricos no teste inicial, mas foram encontrados limites categóricos baseados em duração ou qualidade no segundo teste.

Cebriani (2002), para investigar a percepção do contraste tenso/frouxo das vogais inglesas por falantes nativos de catalão, utilizou um contínuo /i:/i/ε/, variando em qualidade onze vezes; e em duração quatro vezes. Nesse estudo, também a duração foi considerada o parâmetro mais utilizado pelos falantes não nativos na percepção deste contraste.

Fox e Maeda (1999) utilizaram vogais com a duração manipulada para medir os parâmetros adotados por japoneses falantes de inglês como LE na percepção do contraste tenso/frouxo entre as vogais inglesas /i/ e /i/. Nessa pesquisa, foi feito um pré-teste e, em seguida, duas semanas de treinamento, nas quais um grupo de sujeitos treinou com estímulos naturais; e outro, com estímulos que tiveram informações relativas à duração removida. Em seguida, foi feito um pós-teste para medir a eficiência desse treinamento. No pré-teste, pôde-se observar que os sujeitos utilizaram a duração como principal parâmetro na identificação da vogal. Ambos os grupos obtiveram melhores resultados no pós-teste, mas ficou claro que o treino, focando-se apenas na qualidade da vogal, pode ser tão ou mais eficiente que treinamento utilizando qualidade e duração.

Vogais sintetizadas têm sido utilizadas também para comparar a produção e percepção da LE. Flege, Bohn e Jang (1997) pesquisaram

o efeito da experiência com a língua inglesa na produção e percepção das vogais dessa língua. Foram utilizados para esse teste, 80 falantes de inglês como LE e 10 falantes como LM. Dois contínuos de vogais sintéticas foram produzidos: um de *beat* /i/ para *bit* /ɪ/, e outro de *bat* /æ/ para *bet* /ɛ/. Em ambos os contínuos, a qualidade foi modificada 11 vezes e a duração 3 vezes, produzindo um total de 66 estímulos, apresentados depois em um teste de identificação. Os dados obtidos a partir desses testes foram comparados com os de um teste de produção também elaborado. Pôde-se observar que os falantes não nativos com mais experiência na LE produziam e percebiam as vogais inglesas melhor que os inexperientes. Também foi apontada uma relação entre o grau de precisão da produção com o de percepção das vogais inglesas. Finalmente, pôde ser visto que a precisão na produção e percepção dessas vogais variavam em relação ao inventário fonético da LM.

A relação entre produção e percepção também foi examinada por Rochet (1995) que solicitou a falantes canadenses de inglês e falantes brasileiros de português que imitassem a vogal francesa /y/ e catalogassem um contínuo de vogais altas como /i/ ou /u/. Os erros de produção – a substituição da vogal francesa /y/ por /u/ no caso dos falantes de inglês, e por /i/ no caso dos de português – correspondiam aos resultados no teste de percepção – vogais com o segundo formante na altura de /y/ eram mais identificadas como /u/ pelos falantes de inglês e como /i/ pelos de português.

Nessas seis pesquisas, o uso de estímulos sintéticos foi apontado como eficiente, todavia alguns detalhes merecem revisão. Somente Flege, Bohn e Jang (1997) e Morrison (2002) apresentaram as vogais dentro de um núcleo silábico (/bVt/-/bVd/) e somente Morrison as apresentou em uma sentença de suporte. As demais pesquisas apresentaram as vogais isoladas, o que não garante uma percepção fonêmica. Também importante, é que embora haja evidências da utilização da ditongação na discriminação das vogais inglesas (NEAREY, 1989), nenhum dos estudos revisados levou esse parâmetro em consideração. Por fim, todos esses estudos investigaram um número pequeno de vogais, tratando a sua aquisição como a distinção

entre um contraste tenso/frouxo. Assim, em todos eles é feita referência à qualidade espectral das vogais, mas não é feita relação das vogais com um *sistema vocálico*, apesar da existência de estudos que apontem que é, em relação com esse sistema, que a qualidade da vogal é percebida (LADEFOGED, 1989; LADEFOGED & BROADBENT, 1957) e de que o aprendiz de uma língua estrangeira não constrói categorias fonéticas apenas das vogais individuais, mas também do sistema vocálico ao qual pertencem, com o posicionamento de cada vogal no espaço acústico relativo à posição das demais (BAPTISTA, 1992a, 1992b, 2000a, 2000b, 2002).

3. PRESENTE ESTUDO

Objetivando o desenvolvimento de um teste que verificasse os limites categóricos na percepção de *todas* as vogais anteriores do inglês, incluindo o semiditongo /ei/, e que representasse com segurança a criação, ou não, de uma nova categoria fonética, elaborou-se o presente teste. Este teste, cujos resultados serão posteriormente comparados com os de um teste de produção envolvendo os mesmos sujeitos, possui como objetivo maior avaliar a interligação das vogais dentro de um sistema integrado, fornecendo dados de percepção que indiquem a aquisição da supercategoria fonética do sistema vocálico integral da LE. Para tanto, será verificado se um limite categórico muito alto/baixo entre um par de vogais está ligado com a falta de percepção da distinção entre outro par. Entretanto, tal análise não cabe ao presente estudo que se concentrará em esclarecer a metodologia do teste e em avaliar suas potencialidades.

3.1. Sujeitos

Testou-se um total de 25 participantes brasileiros (5 homens e 20 mulheres), com idades entre 18 e 53 anos, e 3 americanos (2 homens e 1 mulher), com idades de 26, 43 e 55 anos. Os participantes brasileiros eram estudantes do curso de *Letras Inglês* da *Universidade Federal de Santa Catarina* e estavam cursando a disciplina *Fonética*

e Fonologia do Inglês. Eles possuíam nível intermediário/avançado de proficiência e já haviam tido alguma espécie de treinamento de pronúncia, o que deveria resultar em limites categóricos bem definidos. Todos os sujeitos declararam não ser portadores de nenhuma deficiência auditiva ou da fala na LM.

3.2. Estímulos

Uma foneticista, falante nativa de inglês, leu 5 sentenças ('I said bVt now') contendo todas as vogais anteriores do inglês, incluindo o semiditongo /eɪ/, em uma sala com isolamento acústico em um estúdio profissional de gravação. Essas sentenças foram gravadas com taxa de amostragem de 44.100Hz, 16-bits, e depois transferidas para um CD de dados.

No *Laboratório de Fonética Acústica da UFSC*, as sentenças tiveram sua taxa de amostragem reduzida (com filtro *anti-alias*) para 7.000Hz. Dessa maneira, a frequência Nyquist ficou em 3.500Hz, pouco acima da maior frequência a ser modificada (3º formante de *beat*: 3.259Hz).

Em seguida, as palavras foram extraídas das sentenças de suporte e tiveram suas vogais normalizadas por pico de intensidade com o auxílio do programa *Sound Forge 6.0*. Isso feito, as 5 palavras foram carregadas no *Laboratório de Análise e Síntese (ASL)* da *Kay Elemetrics Corp*. Nesse programa, cada pulso glotal da vogal foi manualmente marcado, para que pudesse ser feita a análise LPC de cada quadro. Os resultados desses quadros individuais são exibidos pelo programa como uma série de valores numéricos, com cada fileira horizontal representando um pulso glotal a ser analisado ou modificado para síntese. No presente teste, os parâmetros modificados foram (1) duração e (2) frequências formânticas.

- (1) A duração das vogais foi igualada de acordo com o membro mais longo dos grupos /i/(0,21s.),/ɪ/(0,17s.) e /eɪ/(0,23s.); e /æ/(0,20s.)-/ɛ/(0,22s.). Com a duração igualada, os sujeitos possuem informações relativas apenas à qualidade das vogais para identificar os estímulos.

- (2) Os três primeiros formantes das vogais foram modificados em contínuos com sete membros igualmente espaçados na escala mel – uma escala auditiva, e não física como a escala Hz (HAYWARD, 2000; LADEFOGED, 1996). Os valores dos formantes de cada membro do contínuo foram calculados através das médias entre os valores dos pulsos glotais individuais das duas vogais originais de cada par. Dessa maneira, a ditongação das vogais foi também gradualmente modificada, outra inovação desse teste. Isto pode ser exemplificado pela tabela 1, na qual as colunas verticais reproduzem os valores originais de 10 pulsos glotais de *beat*, de *bait* e os valores calculados para os respectivos pulsos glotais dos 7 membros do contínuo.

Tabela 1

	bait	beat	1	2	3	4	5	6	7
1	416	280	280	302	324	346	369	392	416
2	407	275	275	296	318	339	362	384	407
3	402	274	274	294	315	336	358	380	402
4	397	271	271	291	312	333	354	375	397
5	381	271	271	289	307	325	343	362	381
6	380	271	271	289	306	324	343	361	380
7	363	271	271	286	301	316	332	347	363
8	367	271	271	287	302	318	334	351	367
9	367	270	270	286	302	318	334	350	367
10	359	269	269	284	298	313	328	344	359

Nesta pesquisa, ao contrário do que foi feito nas precedentes, os pares foram sintetizados em ambas as direções, ou seja: de *beat* /bit/ para *bit* /bit/ e de *bit* /bit/ para *beat* /bit/, de *beat* /bit/ para *bait* /bet/ e de *bait* /bet/ para *beat* /bit/, de *bit* /bit/ para *bait* /bet/ e de *bait* /bet/ para *bit* /bit/, e por fim, de *bat* /bæt/ para *bet* /bet/ e de *bet* /bet/ para *bat* /bæt/. Para os limites categóricos serem tomados como verdadeiros, eles deveriam apresentar-se iguais, ou muito próximos, nos dois contínuos de cada par.

Finalmente, as palavras foram recolocadas dentro da sentença de suporte original e gravadas em um CD de áudio em ordem aleatória, com um intervalo de 0,7s. entre cada sentença dos pares e de 1,5s. entre um par e outro. Os estímulos foram precedidos de uma gravação correspondente ao seu respectivo número na folha de resposta e agrupados em blocos de 10, evitando-se, desse modo, que os ouvintes cometessem erros em suas marcações.

3.3. Procedimentos

O CD gravado foi tocado no aparelho Phillips CDR-570 no *Laboratório de áudio do curso de Letras da UFSC*, onde os alunos ouviram o som através de fones de ouvido em mesas individuais. Foi fornecida uma breve instrução informando aos sujeitos o procedimento do teste: eles ouviriam 61 pares de sentenças e marcariam na folha de resposta se ouviam a mesma vogal na palavra central.

Foi feito, como aquecimento, um pré-teste com 5 pares de sentenças, após o qual, os participantes declararam não possuírem nenhuma dúvida. Pôde-se, então, tocar os 56 pares (7 para cada um dos 8 contínuos) em um teste que durou 12 min.

4. RESULTADOS

As duas diferentes combinações de cada par de vogais continham elementos com a mesma frequência formântica e duração. Assim sendo, poder-se-ia, no par de *beat /bit/* para */bit/*, por exemplo, tomar como verdadeiro que, quando o sujeito marcava o par como igual, estava identificando o segundo membro do contínuo como sendo */i/*, da mesma maneira, quando o marcava como diferente, o estava identificando como */I/*. O inverso esperava-se do contínuo variando de *bit /bit/* para *beat /bit/*. Da média entre as porcentagens de identificação dos estímulos como uma determinada vogal nos dois contínuos, foram calculados os limites categóricos na percepção dos pares de vogais. Caso uma grande diferença fosse observada entre os dois membros tomados como iguais, não se poderiam considerar os limites categóricos como verdadeiros.

Em alguns contínuos, puderam ser observados limites categóricos muito bem definidos tanto pelos falantes nativos, como pelos aprendizes brasileiros. Este foi o caso das vogais /i/ e /e/ que mostraram, nos dois contínuos, respostas muito próximas. O gráfico 1 traz as porcentagens de identificação da vogal /i/ nos contínuos de *beat* /bit/ para *bit* /bit/, de *bit* /bit/ para *beat* /bit/ e suas médias. O gráfico 2 reproduz as porcentagens de identificação da vogal /e/ nos contínuos de *beat* /bit/ para *bait* /beit/, de *bait* /beit/ para *beat* /bit/ e novamente suas médias. Fica claro, nos dois gráficos, que os limites mostraram-se estáveis em ambas as direções.

Gráfico 1

percepção de /i/ por falantes não nativos

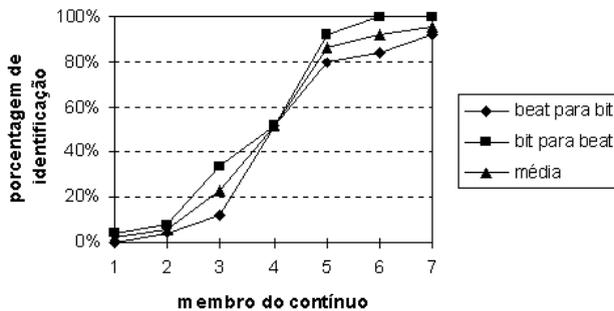
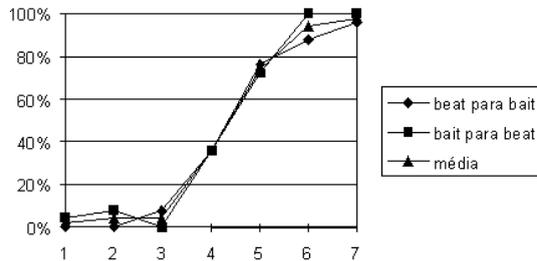


Gráfico 2

percepção de /e/ por falantes não nativos



Entre outros pares de vogais, não puderam ser desenhados limites categóricos seguros. Este foi o caso das vogais /æ/-/ɛ/, e /ɪ/-/eɪ/. Nesses pares não houve correspondência entre os limites obtidos nos dois contínuos. Esse limite pouco definido ocorreu possivelmente devido à qualidade da síntese – alguns desses estímulos podiam ser reconhecidos como modificados, fato inclusive apontado por um dos sujeitos. Isso teria causado uma estranheza por parte dos participantes, que identificariam estes pares como diferentes por terem reconhecido um estímulo sintético, e não por terem reconhecido vogais distintas. Uma evidência para essa hipótese é que mesmo os falantes nativos não apresentaram respostas coerentes nesses pares. Outra explicação pode estar no fato de cada estímulo ter sido tocado apenas uma vez, não havendo assim confirmações das respostas dos sujeitos.

5. CONCLUSÕES E IMPLICAÇÕES

Neste trabalho, discorreu-se sobre o uso de estímulos sintéticos em testes de percepção de vogais de uma LE e avaliaram-se as potencialidades de um novo teste de percepção. Esse teste propôs representar com segurança a criação de novas categorias fonéticas e delimitar os limites categóricos na percepção das vogais anteriores do inglês.

A metodologia do teste mostrou-se eficiente, podendo servir de modelo para os testes posteriores. O uso de estímulos produzidos através da edição de fala naturalmente produzida proporcionou a possibilidade de apresentar as vogais dentro de palavras, e até mesmo de sentenças de suporte, garantindo assim uma percepção fonêmica, e não auditiva. A criação de dois contínuos para cada par de vogal serviu como segurança na verificação de seus limites categóricos. O fato de a duração das vogais ter sido igualada garantiu que os ouvintes estavam baseando suas respostas apenas na qualidade espectral das vogais. Outra importante inovação, foi a elaboração de um contínuo no qual a ditongação das vogais também foi gradualmente modificada.

Embora o teste tenha se mostrado eficiente, alguns detalhes em sua elaboração necessitam ser revistos antes de sua aplicação

definitiva. A qualidade da síntese poderia ser aperfeiçoada, pois, em alguns casos, as modificações nos formantes causaram uma perda de naturalidade que não foi possível recuperar. Outra necessidade confirmada na análise dos resultados foi a de tocar-se, pelo menos três vezes, cada par de estímulos para garantir uma maior confiabilidade dos resultados.

NOTA

¹ O presente trabalho foi realizado com o apoio do *Conselho Nacional de Desenvolvimento Científico e Tecnológico* – CNPq – Brasil. Meus agradecimentos à Profa. Dra. Barbara O. Baptista pelo seu apoio durante esta pesquisa e por suas sugestões em uma versão preliminar deste trabalho. Qualquer correspondência deve ser endereçada para ricardobion@hotmail.com.

REFERÊNCIAS

BAPTISTA, B. O. *The acquisition of English vowels by eleven Brazilian Portuguese speakers: An acoustic analysis*. 1992a. Tese de doutorado - University of California, Los Angeles, EUA.

BAPTISTA, B. O. Sufficient perceptual distance and the acquisition of English vowels. In: MEETING OF THE AMERICAN ASSOCIATION FOR APPLIED LINGUISTICS, 1992b, Seattle, WA, EUA.

BAPTISTA, B. O. A aprendizagem fonética do sistema vocálico de uma língua estrangeira. In: CONGRESSO NACIONAL DE FONÉTICA E FONOLOGIA, 2000, Universidade Federal Fluminense, Niterói, RJ, 2000a.

BAPTISTA, B. O. *The acquisition of English vowels by Brazilian-Portuguese speakers*. Florianópolis, SC: Universidade Federal de Santa Catarina, 2000b.

BAPTISTA, B. O. Adult phonetic learning of a second language vowel system. In: NEW SOUNDS 2000, 2000, Amsterdam: Universidade de Amsterdam. *New Sounds 2000: Proceedings of the Third International Symposium on the acquisition of second-language speech*. Klagenfurt: University of Klagenfurt, 2002. p. 32-41.

BEST, C. T. A direct realist view of cross-language speech perception. In: STRANGE, Winfred (Org.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Timonium, MD: York, 1995. p. 171-203.

CARAMAZZA, A., YENI-KOMSHIAN, G. H., ZURIF, E. B., CARBONE, E. The acquisition of a new phonological contrast: the case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America*, v. 54, p. 421-428, 1973.

CEBRIAN, J. Acquiring a new vowel contrast: The perception of English lax-tense vowels by native Catalan subjects. In: JAMES, A.; LEATHER, J. (Ed.). *Proceedings of the "New Sounds 2000"*. 4th International Symposium on the Acquisition of Second Language Speech. Amsterdam: Universidade de Amsterdam, 2002. p. 48-57.

ESCUADERO, P. *Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels in Spanish speakers*. 2000. Dissertação (Mestrado) - Universidade de Edimburgo.

ESCUADERO, P. The perception of English vowel contrasts: Acoustic cue reliance in the perception of new contrasts. In: JAMES, A.; LEATHER, J. (Ed.). *Proceedings of the "New Sounds 2000"*. 4th International Symposium on the Acquisition of Second Language Speech. Amsterdam: Universidade de Amsterdam, 2002. p. 122-131.

FLEGE, J. E. Second language speech learning theory, findings and problems. In: STRANGE, Winfred (Org.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Timonium, MD: York, 1995. p. 233-277.

FLEGE, J. E. Assessing non-natives' perception of English vowels: A categorial discrimination test. Manuscrito não publicado. Birmingham, AL: University of Alabama, [s.d.].

FLEGE, J. E.; BOHN, O. S.; JANG, S. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, v. 25, p. 437-470, 1997.

FLEGE, J. E.; EFTING, W. Imitation of a VOT continuum by native speakers of English and Spanish: evidence for phonetic category formation. *Journal of the Acoustical Society of America*, v. 83, p. 729-740, 1988.

FOX, M.; MAEDA, K. (1999). Categorization of American English vowels by Japanese Speakers, [Internet]. *Proc. ICObS99*. p. 1437, 1440. [2002].

HAYWARD, K. *Experimental Phonetics*. Padstow, Cornwall: T.J. International Ltd, 2000. p. 140.

LADEFOGED, P. A note on 'Information conveyed by vowels'. *Journal of the Acoustical Society of America*, v. 85, p. 2223-2224, 1989.

LADEFOGED, P. *Elements of Acoustic Phonetics*. 2.ed. Chicago: The University of Chicago Press, 1996. p. 80.

LADEFOGED, P.; BROADBENT, D. Information conveyed by vowels. *Journal of the Acoustical Society of America*, v. 29, p. 98-104, 1957.

MORRISON, G. S. *Effects of L1 duration experience on Japanese and Spanish Listeners' Perception of English High Front Vowels*. Unpublished master's thesis. Simon Fraser University. Burnaby, BC. Canada, 2002.

MORRISON, G. S. Perception of English /i/ and /ɪ/ by Japanese and Spanish Listeners: Longitudinal Results. In: MORRISON, G. S.; ZSOLDOS, L. (Ed.). *Proc. of the North West Linguistics Conference 2002*. Burnaby, BC, Canada: Simon Fraser University Linguistics Graduate Student Association, 2002. p. 29-48.

NEAREY, M. T. Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America*, v. 85, p. 2088-2113, 1989.

ROCHET, B. L. Perception and production of second-language speech sounds by adults. In: STRANGE, Winfred (Org.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Timonium, MD: York, 1995. p. 379-411.

Para a melhoria da síntese articulatória das vogais nasais do português europeu: estudo da duração e de características relacionadas com a fonte glotal

António Teixeira
Francisco Vaz
Lurdes Castro Moutinho
Rosa Lídia Coimbra
Raquel Castro Lisboa
Universidade de Aveiro - Portugal

Abstract

The aim of this research was to continue the work in articulatory synthesis. Thus, a database was collected, containing the nasal sounds of European Portuguese (EP). This corpus includes the different phonetic contexts for the nasal vowels of EP, focusing on the regional variants. The recordings include voice signal, and, simultaneously, information about glottal source, obtained through an Electroglotograph (EGG).

After the signal processing the following parameters were analysed: duration, F0, open quotient, jitter and shimmer. Finally, statistical and inferential analysis was made using the programme SPSS.

I. INTRODUÇÃO

Motivações

Para a continuidade do trabalho em síntese articulatória na Universidade de Aveiro (TEIXEIRA, 2000), tornam-se imprescindíveis dados acerca da produção dos sons do português europeu. Tendo-se a pesquisa, até ao momento, centrado no estudo das vogais nasais, após a recolha de dados usando EMMA (TEIXEIRA & VAZ, 2001), encontra-se em fase de recolha e análise um novo *corpus* acústico (projeto “Fonética Aplicada ao Processamento da Fala: As Nasais do Português”, <http://www.ii.ua.pt/cidlc/gcl/default.asp>).

O sintetizador articulatório necessita, para além da configuração dos articuladores fora do âmbito desta apresentação, de informação acerca da duração e de parâmetros relacionados com a excitação glotal dos quais se destacam a frequência fundamental (F0) e quociente de abertura (OQ). Uma lista completa dos parâmetros utilizados pelo sintetizador relativos à fonte encontra-se na Tabela 1.

Tabela 1 – Parâmetros relativos à fonte glotal utilizados pelo sintetizador articulatório da Universidade de Aveiro. Em negrito apresentam-se os parâmetros estudados no presente trabalho. Refira-se que alguns dos parâmetros não são facilmente obtidos através de análise do sinal de voz e/ou EGG

<i>Parâmetro</i>	<i>Descrição</i>	<i>Mínimo</i>	<i>Máximo</i>	<i>Valor típico</i>	<i>Unidades</i>
ag0	Área glotal mínima	0	10	0	cm ²
agmax	Área glotal máxima	0	10	0.3	cm ²
f0	Freq. fundamental	20	1000	100-200	Hz
openq	Quociente abertura	0	100	60	% of T0
speed	Quociente Velocidade	0	100	2	
jitter	Variação de F0	0	10	2	%
shimmer	Variação de Agmax	0	10	5	%
slope	Declive	0	100	0.03	cm
aspiration	Aspiração	0	5		
lungs	Pressão pulmonar	0	20000	10000	dyne/cm ²
g0	Ganho	0	100		

Geralmente, os *corpora* existentes e as análises publicadas sobre este tema limitam-se a um número reduzido de falantes, regiões do país e contextos fonéticos.

Projeto “Fonética Aplicada ao Processamento da Fala: As Nasais do Português”

Este trabalho é motivado pela existência de uma investigação multidisciplinar, entre lingüistas e engenheiros (Departamento de Línguas e Culturas – CLC e Departamento de Eletrônica e Telecomunicações – IEETA da Universidade de Aveiro), com o objetivo de estudar e aplicar as técnicas da síntese articulatória ao estudo e síntese do PE.

Desde 1995 que se tem desenvolvido pesquisa no domínio da síntese dos sons nasais do português (TEIXEIRA, 2001). As experiências efetuadas mostram que a variação do velo, e mesmo de outros articuladores, influencia a produção e percepção da nasalidade. No entanto, não existe um estudo acústico detalhado, considerando a variação das características ao longo da produção das vogais nasais. Informação acerca do comportamento da fonte glotal durante a produção dessas vogais é também necessária para a continuação deste trabalho, bem como informação acerca da sua variação regional.

Pelo exposto, constata-se que se torna imprescindível a obtenção de dados sobre a forma como esses sons são naturalmente produzidos. Foi já recolhida informação acerca da posição da língua, lábios e véu palatino, durante a produção de palavras e frases contendo nasais (TEIXEIRA & VAZ, 2001a e 2001b), utilizando um sistema de *ElectroMagnetic Midsagittal Articulography* (EMMA). Esta técnica, no entanto, não é viável para um número alargado de falantes nem fornece informação acerca do processo de fonação.

Assim, neste novo projeto – “Fonética Aplicada ao Processamento da Fala: As Nasais do Português” (<http://www.ii.ua.pt/cidlc/gcl/fap/>) – utilizamos o Electroglotógrafo (EGG) para obter informações acerca da excitação glotal ocorrida no momento de produção das nasais do

português; frequência fundamental das diferentes vogais; quociente de abertura e de velocidade. Uma análise baseada unicamente no sinal de fala impossibilita a utilização de métodos que permitam a filtragem do fenômeno de inversão glotal. O EGG, de simples utilização e não interferindo com o sinal de fala, revela-se um instrumento de grande utilidade numa pesquisa deste tipo e indispensável para a análise da variação sincrónica do pitch.

II. NOVO CORPUS

O *corpus* (MOUTINHO et al., 2001) foi organizado de forma a contemplar diferentes contextos fonéticos e a coleta inclui já diversas variantes regionais: Minho, Douro Litoral, Beiras, Alentejo e Algarve. Para além do sinal de voz, foi recolhido sinal de um electroglotógrafo (EGG), para facilitar a obtenção de parâmetros relacionados com a fonte.

Contextos fonéticos

Na constituição do novo *corpus*, teve-se a preocupação de incluir todas as vogais nasais do português, inseridas nos diversos contextos fonéticos em que podem ocorrer:

- consoante oclusiva surda – vogal nasal – consoante oclusiva surda;
- consoante oclusiva sonora – vogal nasal – consoante oclusiva sonora;
- consoante oclusiva surda – vogal nasal – consoante oclusiva sonora;
- consoante oclusiva sonora – vogal nasal – consoante oclusiva surda;
- consoante fricativa – vogal nasal – consoante fricativa;
- outra consoante – vogal nasal – consoante fricativa;
- consoante fricativa – vogal nasal – outra consoante;
- vogal nasal seguida ou precedida de consoante líquida;
- vogal nasal no início e no fim de palavra [*corpus* lido];
- seqüências (incluindo vogal nasal) não existentes dentro de palavra;
- vogal oral – vogal nasal.

Foram também contemplados, no *corpus*, contextos para as vogais orais, com o objetivo de confrontar, nos mesmos falantes e nos mesmos contextos, as características das vogais nasais estudadas com as orais correspondentes.

Este *corpus* consideramos vir colmatar uma lacuna no conhecimento dos sons nasais do PE, pelo fato de contemplar variantes regionais e fornecer informações, pela utilização do EGG, acerca da fonte glotal, até ao momento, não considerada.

Cobertura Regional

A análise que aqui será apresentada fundamentar-se-á nos resultados referentes a seis informantes do sexo masculino provenientes de três das regiões já disponíveis: Beira Litoral, Alentejo e Algarve.

A escolha recaiu sobre informantes com idades superiores a 35 anos, naturais e residentes nas localidades escolhidas, não tendo uma escolaridade superior ao atual ensino obrigatório ou seu correspondente. Esta preocupação prende-se com o fato de se pretender obter uma amostragem homogênea, de modo a apenas fazer variar o parâmetro geográfico. Note-se, ainda, que este perfil de informante deverá garantir-nos uma maior representatividade da respectiva variante geográfica.



Figura 1 - Mapa de Portugal com indicação das zonas onde foi efetuada a coleta do material analisado neste artigo.

Recolha



Figura 2 - Fotografia de uma sessão de gravação. São visíveis o sistema de aquisição e o computador portátil (*notebook*)

As gravações foram sempre realizadas no terreno, utilizando o seguinte equipamento: um CSL 4400 e EGG 6103, ambos da Kay Elemetrics Corp., e um computador portátil, tendo o sinal sido captado através de um microfone Shure SM48 diretamente para o disco duro. Durante a coleta, foram utilizados, sempre que possível, estímulos visuais, através de imagens que conduzissem o informante a produzir as palavras pretendidas, evitando, assim, a sua leitura. A cada informante foram solicitadas duas repetições do *corpus*.

Software utilizado

Para este projeto escolhemos o sistema SFS desenvolvido Mark Huckvale do University College London que pode ser obtido a partir de <http://www.phon.ucl.ac.uk/resource/sfs>.

Tanto para a gravação, como para a segmentação, foi utilizado o programa SFSWin, o qual possibilita a gravação simultânea dos dois canais de entrada que se pretendem utilizar: o microfone e o electroglotógrafo. O SFSWin permite visualizar todas as formas de onda, em simultâneo ou individualmente, (oscilograma, espectrograma e

o sinal acústico que representa os movimentos das cordas vocais na produção do som). Para além disso, permite também operar sobre determinados segmentos selecionados, ouvir, apagar, anotar, etc. O sistema SFS permite a conversão dos arquivos no formato SFS em arquivos WAV, de forma a poderem ser processados por outros programas, nomeadamente pelo MatLab. O programa possui uma ferramenta que permite a visualização das formas de onda, em tempo real, do sinal acústico. Assim, é possível monitorizar a forma de onda e a forma espectral dos sons que estão sendo gravados.

Segmentação e Etiquetagem

Após gravação, procedeu-se a uma primeira audição, com vista a selecionar as palavras que constituirão os arquivos a analisar. Para cada um dos segmentos correspondentes às palavras do *corpus* é criada uma anotação no seu início e fim. Com a ajuda do comando **wordchop** do SFS são depois criados arquivos individuais. O nome desse arquivo contém informação acerca do informante, tabela e item do *corpus* a que se refere, e qual a repetição. Por exemplo, o arquivo **clrT01I01r1.sfs** contém os dados referentes à primeira repetição do primeiro item da tabela 1 (vogais nasais entre oclusivas surdas) preferido pelo informante CLR.

A segmentação consiste na identificação dos segmentos consonantais que precedem e seguem a vogal nasal, ficando esta, conseqüentemente, delimitada. Esta anotação é realizada utilizando o sistema SAMPA para o português. Apresenta-se, na Figura 3, um exemplo de um arquivo anotado.

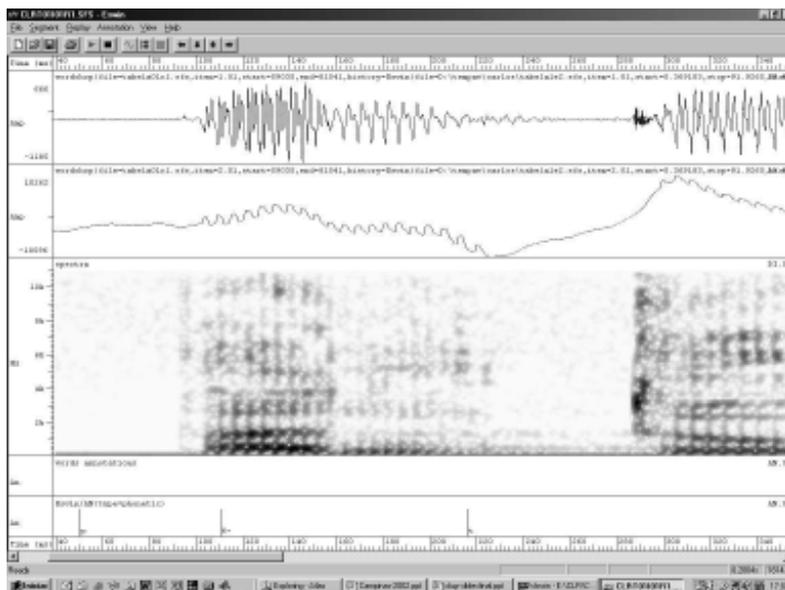


Figura 3 - Exemplo de uma anotação do *corpus*. De cima para baixo, podem ser vistos o sinal de voz, o sinal do EGG, anotação ao nível das palavras e anotação fonética. Nesta última podem ser vistos os pontos de início da vogal nasal e da consoante que a segue (lt) assim como a indicação de qual o segmento que antecede a vogal nasal (neste caso o [p]).

III. METODOLOGIA DE ANÁLISE

- Procura de uma forma automática, em todos os arquivos, dos segmentos anotados como vogal (oral e nasal) obtenção, para cada um, de: duração, F0, quociente de abertura, *jitter* e *shimmer*. Para cada segmento, informação acerca das medidas obtidas e dos vários fatores (falante, região, sexo do falante, tabela, etc.) é guardada para análise estatística;
- execução de análises estatísticas usando os programas SPSS (Bryman, Cramer 2001) e MacAnova;
- interpretação dos resultados em função da variação regional, das diferentes vogais nos vários contextos fonéticos.

Apresenta-se a seguir alguma informação sobre a forma de obtenção dos parâmetros analisados.

Duração

Como no processo de anotação indica-se onde começa a vogal e onde começa o segmento que a segue. O processo de obtenção da duração é muito simples, envolvendo apenas a utilização do comando **lengthn** do SFS. Como os nossos estudos anteriores apontam para que as vogais nasais sejam constituídas em geral por três fases, constituindo o murmúrio nasal a fase final, não medimos este segmento em separado considerando-o parte integrante da vogal nasal.

Freqüência fundamental (F0)

Tendo sido gravado o sinal EGG, a obtenção de freqüência fundamental foi efetuada de duas formas: uma baseada no EGG, outra usando apenas o sinal de voz .

No primeiro caso, o processo consiste em obter informação acerca do início de cada período do sinal EGG, usando o comando **vtx**, obtendo depois o valor de F0 com base nesta informação usando o comando **fx**. O comando **vtx** estima a posição dos períodos de excitação glotal a partir do sinal do EGG, através do método utilizado no Voiscope, que consiste em: (i) filtragem, (ii) controle de ganho automático e (iii) detecção de picos. Este processo é exemplificado na Figura 4, onde se pode ver o sinal EGG, informação acerca do início e fim de cada período e F0.

No segundo caso, a análise é efetuada usando o comando SFS **fxanal**. Este comando estima F0 do sinal de voz através do método da autocorrelação.

Para possibilitar o estudo de uma possível variação deste parâmetro ao longo da realização da vogal, o valor de F0 foi calculado em três pontos: próximo do início (concretamente a 10% da duração da vogal), no meio e próximo do final (a 90% da duração).

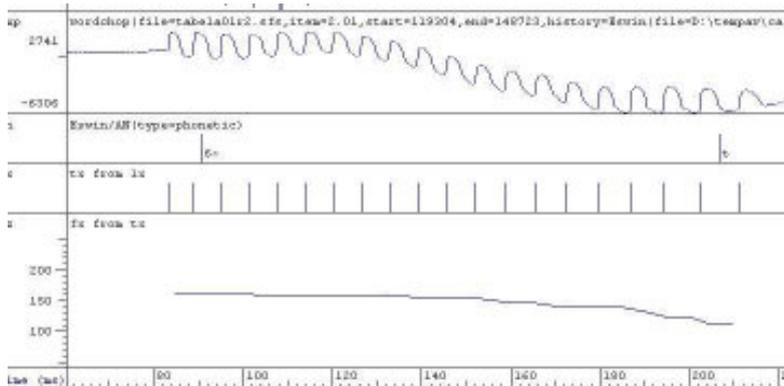


Figura 4 - Exemplo de cálculo de F0 com base no sinal EGG. Com base no sinal EGG, apresentado no topo, é primeiro calculado o início de cada período glotal (chamado de TX no sistema SFS) e depois calculada a frequência fundamental (FX no sistema SFS).

Quociente de abertura

Diversos eventos do processo de fonação estão relacionados com eventos (segmentos ou picos) correspondentes ao sinal EGG ou a sua derivada (DEGG). Por exemplo, o instante de abertura das cordas vocais e o instante do pico positivo no sinal DEGG são coincidentes, assim como o instante de fecho das cordas vocais e o instante do pico negativo do DEGG (CHILDERS, 2000, p. 87). O quociente de abertura (em inglês *open quotient* – OQ) medido da área glotal e o OQ medido a partir do sinal EGG são quase o mesmo, sendo: $OQ = (\text{duração de abertura das cordas vocais}) / (\text{duração do período glotal})$.

O quociente de abertura foi calculado com base na informação acerca da abertura e fecho das cordas vocais gerada pelo comando SFS **txgen**. O processamento efetuado é o seguinte:

- 1) Filtragem passa-baixo do sinal EGG;
- 2) Diferenciação do sinal EGG (obtendo DEGG);
- 3) Procura do máximo e mínimo do sinal DEGG que se encontra acima de um determinado limiar. A posição destes é utilizada como posição de fecho e início de abertura.

O resultado desse processamento é exemplificado na Figura 11. Com base nesta informação foi desenvolvido um programa em SML (*Speech Measurement Language*) para calcular o OQ. Infelizmente, a detecção efetuada pelo **txgen** apresenta muitas deficiências tornando muito difícil a obtenção de uma boa estimativa de OQ em arquivos com o sinal EGG de menor qualidade.

Jitter

O *Jitter* é uma medida da perturbação do período de oscilação das cordas vocais. Existem várias formas propostas para calculá-lo. No sistema SFS o comando **txstat** permite obter informação do *jitter*. Devido a falta de documentação deste comando, não está disponível a informação acerca do método utilizado. Optamos por desenvolver um novo programa de cálculo do *jitter* usando SML optando pelo cálculo do *Jitter Factor* (LALWANI, 1991). O *Jitter Factor* (JF) é definido como a média do valor absoluto das perturbações na frequência fundamental, normalizada pela média do valor da frequência fundamental, ou seja:

$$JF = \frac{\frac{1}{N-1} \sum_{i=2}^N |F_{0_i} - F_{0_{i-1}}|}{\frac{1}{N} \sum_{i=1}^N F_{0_i}} = \text{Média do módulo das variações} / \text{Média de F0}$$

O valor do JF é calculado usando uma janela (de comprimento N). Para permitir estudar possíveis variações ao longo da vogal, o JF foi calculado: (i) considerando toda a duração da vogal, (ii) uma janela de comprimento fixo (N=5) no início, (iii) no meio e (iv) no fim. Como foi notado, em muitas realizações das vogais, um decréscimo do valor de F0 ao longo da vogal, o valor do *jitter* devido à utilização de um menor valor de F0 média na normalização terá tendência para atingir valores mais elevados. Por este motivo, para além de guardar o JF calculado das 4 formas anteriormente referidas, foi guardado para análise o valor do numerador da fórmula de cálculo do JF.

IV. RESULTADOS

Duração

A primeira análise com interesse para aplicação em síntese consiste em obter os valores médios das durações. Na tabela 2, apresentam-se os valores médios, utilizando 6 falantes de 3 regiões, para as 8 vogais orais e as 5 vogais nasais. Os resultados foram decompostos tendo em conta se a vogal se encontrava acentuada ou não.

Tabela 2 – Durações médias das vogais analisadas.

	oral		nasal	
	não acentuada	acentuada	não acentuada	acentuada
a	,	109,38	,	,
Λ	65,81	121,98	133,94	159,14
e	,	112,03	147,49	169,12
ε	,	105,97	,	,
i	,	104,78	147,35	170,76
o	,	116,2	,	168,55
ɔ	83,28	111,0	,	,
u	66,56	126,1	,	173,66

É notória a contribuição da nasa para a duração. As vogais orais não acentuadas apre as vogais nasais em posição acentuada, as maiores durações.

Para além dos valores médios, interessa também a dispersão em torno desses mesmos valores. Na Figura 5 apresenta-se o intervalo de confiança a 95% da média.

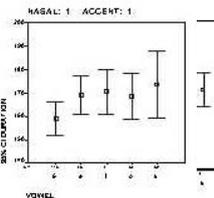
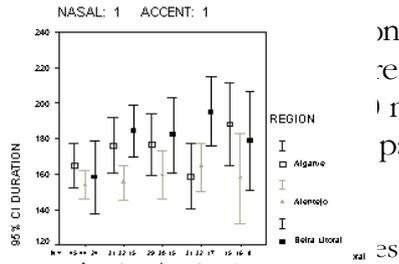


Figura 5 – Intervalo de confiança a 95% e valor médio da duração das vogais orais e nasais. Os dois gráficos da esquerda referem-se a vogais orais (nasal=0), o primeiro e terceiro gráfico referem-se ao caso não acentuado (accent=0).

Será que este comportamento se verifica para as três regiões? Para tentar obter uma resposta a esta questão foi efetuada a análise cujos resultados se apresentam na Figura 6. Como estamos mais interessados nas nasais, apenas se apresentam os gráficos referentes a estas vogais.

Figura 6 – Intervalo de confiança das durações para as vogais nasais em posição não acentuada (esquerda) e acentuada (direita) apresentando separadamente os valores obtidos para cada região.

Apesar das médias serem algo sobrepostas. No caso não acentuado de 100 e 180 ms, no caso acentuado Na Tabela 3 apresentam-se três regiões analisadas.



Intervalo de confiança de cerca de 100 ms. para as

Tabela 3 – Médias de duração

		VOWEL		Litoral
		NON-ACCENTUATED	ACCENTUATED	
A	não acentuada	146,03	125,49	132,68
	acentuada	164,80	153,80	158,33
E	não acentuada	161,65	130,69	151,35
	acentuada	176,17	155,14	184,36
I	não acentuada	153,22	135,37	148,80
	acentuada	176,59	159,43	182,15
O	não acentuada	141,86	121,25	147,08
	acentuada	158,75	163,93	195,10
U	não acentuada	151,03	132,92	135,87
	acentuada	187,91	157,85	178,57

Para possibilitar a comparação com o português do Brasil, apresenta-se, na Tabela 4, um resumo das durações médias do presente estudo e do estudo realizado por Sousa (1994). Comparando os valores apresentados nas duas últimas colunas, constatamos não serem detectáveis diferenças significativas entre os valores obtidos para estas duas variedades do português.

Tabela 4 – Comparação dos valores médios obtidos com os de Sousa (1994) para o português do Brasil. As duas primeiras colunas de valores apresentam os resultados de Sousa; a última, os resultados obtidos no presente estudo.

Vogal	de Sousa (1994)		Nosso resultado Para nasal
	Oral	Nasal	
[a] x [ã]	123,03	163,54	159,14
[e] x [ê]	109,77	170,05	169,12
[e] x [ɛ]	98,17		
[i] x [í]	86,85	155,6	170,76
[o] x [ô]	123,2	161,58	168,55
[o] x [ɔ]	106,11		
[u] x [ú]	88,54	144,38	173,66

Embora a influência de outros fatores na duração seja também passível de ser investigada com o material disponível, a ainda não disponibilidade para análise da totalidade do *corpus* levou-nos a adiar essas análises para um futuro próximo.

Os testes estatísticos (análise de variância de medidas repetidas – *repeated measures ANOVA*) confirmam como significativa a influência da vogal [F(4,20)=22.033, p=0.000], nasalidade [F(1,3)=11.971, p=0.041], acento [F(1,3) =10.786, p=0.046 (muito próximo de 0.05)] e posição [F(2,6)=48,702, p=0.000] na duração. A influência da região [F(2,3)=.355, p=0.727 (ns)] não se revelou significativa, assim como a influência da repetição [F(1,3)=8.604, p=0.061 (ns)]. Relativamente a este último resultado, pode considerar-se como um elemento de validação das coletas. Diversas interações entre os fatores revelaram-se também significativas – como por exemplo vogal x nasalidade. Não analisaremos aqui esses resultados.

Frequência Fundamental (F0)

Durante o processo de cálculo, foram monitorizados os valores obtidos pelos dois métodos (usando EGG ou o sinal de voz). Para alguns informantes, o processo baseado no EGG não foi capaz de determinar corretamente o início de cada período, falhando um ou mais inícios ou detectando inícios onde estes não existem. Em consequência dessa deficiente detecção, os valores de F0 apresentam valores ora muito altos ora muito baixos. Para analisar esses resultados foram eliminadas, usando as facilidades do comando **fx**, todas medições de F0 inferiores a 100 e superiores a 200 Hz, visto tratar-se de informantes masculinos e os valores obtidos pela análise do sinal de fala não excederem estes limites.

Começamos por investigar se os dois métodos fornecem valores semelhantes. Na Tabela 5 apresentam-se as correlações entre as medidas efetuadas pelos dois métodos. F0 medido no início e no meio apresentam um valor de correlação próximo de cerca de 0.9 para os dois métodos. Os valores mais baixos de correlação aparecem entre F0 iniciais/médios e médios/finais. Devido aos valores elevados da correlação entre os dois métodos e às falhas do método baseado no EGG, grande parte das análises foram efetuadas utilizando os valores obtidos com base no sinal de voz.

Tabela 5 – Correlações entre as 3 medidas pelos 2 métodos

		FOSTART	FOMED	FOEND	FOSTART2	FOMED2	FOEND2
FOSTART	Pearson Correlation	1	,914**	,660**	,862**	,741**	,672**
	Sig. (2-tailed)	,	,000	,000	,000	,000	,000
	N	901	736	837	901	900	901
FOMED	Pearson Correlation	,914**	1	,617**	,774**	,708**	,586**
	Sig. (2-tailed)	,000	,	,000	,000	,000	,000
	N	736	749	695	749	749	749
FOEND	Pearson Correlation	,660**	,617**	1	,596**	,490**	,837**
	Sig. (2-tailed)	,000	,000	,	,000	,000	,000
	N	837	695	909	909	908	909
FOSTART2	Pearson Correlation	,862**	,774**	,596**	1	,868**	,728**
	Sig. (2-tailed)	,000	,000	,	,000	,000	,000
	N	901	749	909	909	1013	1015
FOSTART2	Pearson Correlation	,741**	,708**	,490**	,868**	1	,607**
	Sig. (2-tailed)	,000	,000	,000	,000	,	,000
	N	900	749	908	1013	1013	1013
FOEND2	Pearson Correlation	,672**	,586**	,837**	,728**	,607**	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	,
	N	901	749	909	1015	1013	1015

** Correlation is significant at the 0.01 level (2-tailed).

Na Figura 7 apresenta-se o estudo da variação de F0 ao longo da vogal em função da vogal, nasalidade e acento. Nota-se um valor de F0 inferior no final da realização das vogais, especialmente notório nas nasais acentuadas.

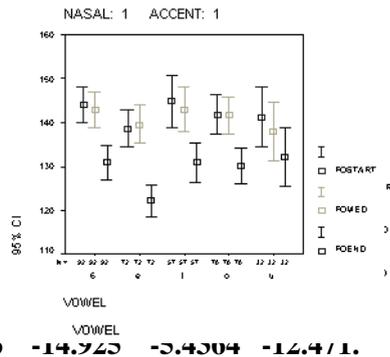
Figura 7 – Intervalo de confiança para nasalidade (nasais em baixo), do acent

Calculando a diferença em vogal, obtém-se:

-15.941 -15.565 -17.272 -16.2

Efetuando o mesmo cálculo

-23.418 -11.851 -18.376



, da reita)

cada

28

1-se:

Ou seja, para todas as vogais e para todos os informantes o valor médio final é cerca de 10 a 20 Hz inferior ao valor médio inicial. O valor médio para vogais orais é -12.854 e -15.02 Hz para as nasais.

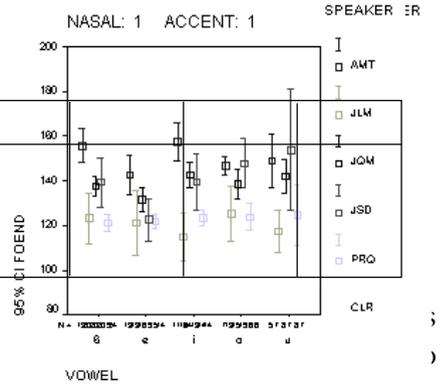
Na Figura 8, apresenta-se o estudo da variação de F0 para os vários informantes. Para a mesma vogal os vários informantes apresentam valores médios e intervalos de confiança díspares. Não consideramos que se possa apresentar um valor médio por vogal. Os fatores intrínsecos ao falante parecem sobrepor-se à vogal. Na Tabela 6

apresentam-se os valores médios no início, meio e fim para cada um dos seis informantes.

Figura 8 – F0 para as vogais nasais (acentuadas ou não) em função da vogal e do falante

Tabela 6 - Valores médios de F0 medido cada um dos sei:

	1 - AMT	2 - JLM	3
F0 inicial	172.81	129.67	160.11
F0 medial	171.91	132.6	160.11
F0 final	149.39	117.82	149.39



Foram efetuadas análises de v no início e fim. Não se justifica a ar pela elevada correlação com o valo

Para o início, o efeito do falante em F0 é estatisticamente significativo, assim como o efeito da posição. Os outros fatores não se revelaram como significativamente afetando o valor de F0 na fase inicial da vogal. Nesta análise, utilizou-se a medida baseada no sinal de fala, para evitar ter de lidar com os casos desconhecidos criados pelas falhas do outro método de obtenção de F0. Detalhes da análise encontram-se na Tabela 7.

Tabela 7 – Resultados da análise de variância (efetuada em Macanova) para os valores de F0 medidos no início

	DF	SS	MS	F	P-value
CONSTANT	2	2.2183e+07	2.2183e+07	28804.00822	0
speaker	5	2.9178e+05	58356	75.77499	0
vowel	7	10657	1522.5	1.97695	0.086576
ERROR1	35	26954	770.12	2.31889	0.17631
nasal	1	8.8699	8.8699	0.02671	0.87658
ERROR2	5	1660.5	332.11	0.19370	0.95207
accent	1	91.851	91.851	0.05357	0.82614
ERROR3	5	8572.9	1714.6	4.86012	0.016292
position	2	4627	2313.5	6.55785	0.015152
ERROR4	10	3527.9	352.79	0.95319	0.48318
ERROR5	943	3.4901e+05	370.11		

Para a medida no final da vogal, além do falante revela-se como também significativa a influência da vogal. A posição onde ocorre a vogal deixou de ser um fator estatisticamente significativo. Detalhes da análise encontram-se na Tabela 8.

Tabela 8 – Resultados da ANOVA realizada para F0 medido no final da vogal.

	DF	SS	MS	F	P-value
CONSTANT	2	1.8089e+07	1.8089e+07	34471.12478	0
speaker	5	1.4281e+05	28563	54.43007	0
vowel	7	15240	2177.2	4.14887	0.020399
ERROR1	35	18367	524.76	0.62862	0.81191
nasal	1	999.09	999.09	1.19614	0.32396
ERROR2	5	4176.3	835.26	0.69276	0.65154
accent	1	549.85	549.85	0.45604	0.52945
ERROR3	5	6028.5	1205.7	1.39928	0.30386
position	2	3078	1539	1.78607	0.21715
ERROR4	10	8616.6	861.66	2.55187	0.004829
ERROR5	943	3.1841e+05	337.66		

Quociente de abertura (OQ)

Na Figura 9 apresenta-se o diagrama de extremos e quartis (*boxplot*) do quociente de abertura em função da vogal e da nasalidade. A figura não revela diferenças de valores médios nem de dispersão para as várias vogais, tanto para as orais como para as nasais.

Figura 9 – Diagrama de extremos e quartis (*boxplot*) dos valores do quociente de abertura em função da vogal e nasalidade.

Na Figura 10 apresenta-se o diagrama de extremos e quartis (*boxplot*) do quociente de abertura para os diferentes informantes. Nota-se uma variação em função do falante, não sendo a região um fator significativo. Falantes da mesma região exibem comportamentos bastante diferentes entre si.

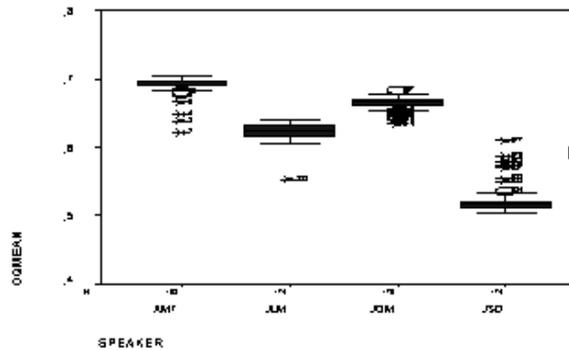


Figura 10 – *Boxplots* para cada falante. Os dois da esquerda do Algarve, os do meio do Alentejo e os da direita da Beira Litoral.

Analisando com mais detalhe, se olharmos para os valores de cada informante para cada uma das vogais, inscritos na Tabela 9, constatamos que todos os valores para o mesmo falante são muito semelhantes, tanto para as orais como para as nasais.

Tabela 9 – Valores médios de OQ por vogal e falante. Na primeira coluna apresenta-se a sigla pela qual é identificado o falante.

oral

	a	Λ	e	ε	i	o	ɔ	u
AMT	,69	,69	,69	,69	,69	,69	,69	,69
JLM	,62	,62	,61	,62	,62	,61	,62	,62
JQM	,67	,66	,67	,67	,67	,67	,67	,67
JSD	,51	,51	,52	,51	,52	,52	,52	,52
PRQ	,62	,62	,58	,58	,58	,58	,58	,58
CLR	,50	,50	,50	,51	,51	,50	,51	,51

nasal

	Λ	e	i	o	u
AMT	,69	,69	,69	,69	,69
JLM	,62	,61	,62	,61	,62
JQM	,66	,67	,67	,67	,67
JSD	,51	,52	,51	,52	,52
PRQ	,62	,58	,58	,58	,58
CLR	,50	,50	,51	,51	,51

Não é utilizável um valor médio global, nem mesmo para cada vogal, visto que cada informante apresenta um valor médio “próprio”. Os histogramas de OQ para cada vogal, em função da região e em função da nasalidade, apresentam vários máximos não sendo representados corretamente por uma distribuição unimodal.

Foi também investigado o efeito do acento, não se tendo constatada uma influência significativa desse fator no quociente de abertura.

Os valores obtidos para quatro dos informante estão de acordo com valores médios publicados para a voz modal (CHILDERS, 2000, p. 313). Os informantes AMT e JQM apresentam valores que já se podem considerar característicos de outros tipos de voz.

Jitter

Apenas foi possível utilizar o método de cálculo do JF em 4 dos 6 informantes. O problema deveu-se à dificuldade que o programa **hqtx** teve em processar o sinal EGG para dois dos falantes.

Na Figura 13 apresentam-se os resultados das medidas efetuadas no início, meio, fim e na totalidade da vogal, em função da vogal e da nasalidade. Numa primeira análise, os valores dos quatro gráficos não diferem grandemente.

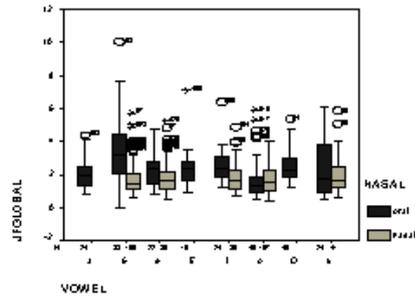


Figura 13 – *Boxplot* do valor do *Jitter* , vogal (início, meio e fim) e usando a totalidade da duração da vogal. São apresentados valores para cada vogal separados pela nasalidade

Os valores médios encontram-se na Tabela 10. Para cada caso apresenta-se o valor de JF no início, meio, fim e na globalidade (de cima para baixo).

Tabela 10 – Valores médios obtidos para o *Jitter Factor*. Na primeira tabela da esquerda por vogal e tendo em conta a nasalidade, a segunda tabela da direita médias globais para os valores medidos nas três posições e na globalidade da vogal. Na última tabela de baixo à esquerda os valores para vogais orais, à direita para vogais nasais.

	a	Λ	e	ε	i	o	ɔ	u
oral	2,16	3,27	2,43	2,82	2,08	1,69	2,42	2,34
	1,53	3,06	1,80	1,81	2,14	1,56	2,26	2,13
	1,96	3,65	2,75	2,36	2,74	1,63	2,20	2,56
	2,07	3,51	2,37	2,47	2,56	1,78	2,59	2,46
nasal	,	1,85	1,88	,	1,91	1,72	,	1,98
	,	1,25	1,51	,	1,28	1,29	,	1,77
	,	2,29	2,17	,	2,47	2,12	,	2,04
	,	1,78	1,85	,	1,83	1,74	,	2,02

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
JFSTART	692	2,0159	,0560
JFMED	692	1,5907	,0526
JFEND	692	2,2996	,0613
JFGLOBAL	692	2,0258	,0433
Valid N (listwise)	692		

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
JFSTART	229	2,3538	,1074
JFMED	229	2,0568	,1062
JFEND	229	2,4284	,1043
JFGLOBAL	229	2,4587	,0972
Valid N (listwise)	229		

a. NASAL = oral

Descriptive Statistics

	N	Mean	
	Statistic	Statistic	Std. Error
JFSTART	463	1,8487	,0560
JFMED	463	1,3602	,0526
JFEND	463	2,2359	,0613
JFGLOBAL	463	1,8116	,0433
Valid N (listwise)	463		

a. NASAL = nasal

Nota-se um valor inferior, relativamente ao início e fim de JF, para as vogais nasais medido a meio da vogal. O valor final é o mais elevado. No entanto, estando o valor de JF dependente do valor médio de F0, é preciso determinar se essa diminuição se deve a um aumento desse valor. Para isso analisamos o valor do numerador da expressão de cálculo de JF, apresentado na Figura 14. Continua a aparecer um valor mais baixo da medida efetuada a meio, pelo menos para 4 das vogais nasais.

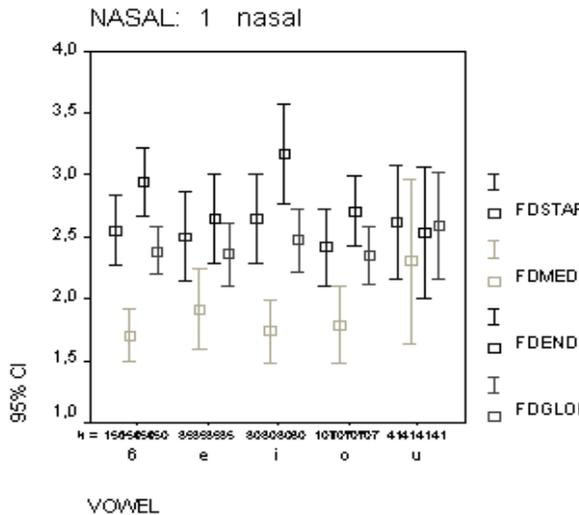


Figura 14 – Numerador do cálculo

Vendo os gráficos e estatísticas associadas com o contexto nasal, o efeito aparece nos dois casos. Este resultado deve ser objeto de mais análises.

A análise de variância revela como não significativa a influência da vogal, nasalidade, acento, posição, repetição e região no valor de JF.

V. DISCUSSÃO

Quanto à variação regional, a amostra apresenta-se ainda demasiado pequena. Por agora, o *corpus* tem apenas como vantagem não ser exclusivamente constituído por falantes de uma zona do país, ser recolhido no terreno, integrando por isso mais variação do que o habitual.

Apesar do *corpus* incluir todos os contextos em que ocorrem vogais nasais em português, neste estudo apenas foram analisados os contextos envolvendo oclusivas, essencialmente por serem os mais representativos e estarem já etiquetados. Os resultados apresentados devem ser lidos tendo em conta a limitação a esses contextos.

A utilização do EGG possibilitou a obtenção de parâmetros como o quociente de abertura e o *jitter*. Na fase de etiquetagem o sinal EGG foi também de grande utilidade, permitindo mais precisão na identificação sobretudo da fase inicial da vogal.

O sistema SFS revelou-se uma ferramenta poderosa, em especial a linguagem SML – *speech measurement language* –, tendo facilitado muito a separação do *corpus* em arquivos individuais e a análise de forma automática de todo o *corpus* já anotado. Foi também fácil criar os arquivos no formato utilizável pelo programa de análise estatística. “Facilmente” se podem criar arquivos para treino de modelos CART para utilização num sistema de síntese. Será interessante analisar as árvores de decisão criadas.

Os resultados apresentados são essencialmente exemplificativos do que se pode obter quando o *corpus* estiver completo. Tencionamos no final do projeto, refazer as análises agora apresentadas utilizando todas as regiões e informantes recolhidos.

Foi feito um esforço de comparação dos resultados obtidos com resultados publicados (ANDRADE, 1987; DELGADO MARTINS, 2002). No entanto, o tipo de parâmetro por nós analisados e a metodologia diferente – tendo em conta os nossos objetivos – adotada para outros não nos permitiu grandes comparações. Por exemplo, muitas das análises de duração das vogais nasais não incluem o segmento consonântico nasal final; muitos dos resultados referem-se apenas a um informante e a voz de “laboratório” não espontânea.

VI. CONCLUSÕES E TRABALHO FUTURO

Motivados pela necessidade de obtenção de parâmetros relativos à fonte glotal para continuação do trabalho que se tem desenvolvido na área de síntese articulatória na Universidade de Aveiro, estamos procedendo a coleta de um *corpus* contemplando as variações contextuais e regionais das vogais nasais. Neste artigo, apresentamos a metodologia de análise e os resultados obtidos relativamente aos seguintes parâmetros: duração, frequência fundamental, quociente de abertura e *jitter*.

Dos resultados obtidos salienta-se:

1. Duração – Os testes estatísticos confirmam como significativa a influência da vogal, nasalidade, acento e posição na duração. A influência da região não se revelou significativa, assim como a influência da repetição. São apresentados: valores médios, utilizando 6 falantes de 3 regiões, para as 8 vogais orais e as 5 vogais nasais; médias para cada vogal para as três regiões analisadas; e a comparação com os valores medidos para o português do Brasil de Sousa (1994);
2. Frequência fundamental – Verifica-se uma grande correlação entre os valores obtidos com base no EGG ou apenas no sinal de voz; um valor de F0 inferior no final da realização das vogais é especialmente notório nas nasais acentuadas; que para a mesma vogal os vários informantes apresentam valores médios e intervalos de confiança díspares, os fatores intrínsecos ao falante parecem sobrepor-se à vogal;
3. Quociente de abertura – Revelou-se uma característica idioletal, cada informante com um valor muito estável com a variação da vogal, nasalidade etc. Falantes da mesma região exibem comportamentos bastante diferentes entre si;
4. *Jitter* - Numa primeira aproximação podem usar-se valores de JF da ordem dos 2%. Caso se esteja interessado em incluir na modelação mais detalhe – cujos efeitos em termos de qualidade deverão ser estudados por testes perceptuais – poderá utilizar-se um valor um pouco inferior a 2% na zona média e um valor acima de 2% na parte final.

Trabalho futuro

1. Analisar todo o *corpus* assim que estiver disponível, estendendo esta análise para todas as regiões do corpus e ao sexo feminino;
2. Melhorar a técnica de obtenção do quociente de abertura;
3. Adicionar a análise da duração da fase final de realização das vogais nasais (o habitualmente designado de murmúrio nasal);
4. Medir outros parâmetros da fonte como quociente de velocidade (*Speed Quotient*);
5. Finalmente, utilizar os dados no sintetizador articulatório da Universidade de Aveiro.

Agradecimentos

A todos os que contribuíram para o *corpus*, FCT pelo financiamento (Projeto PCTI 1999 PLP 36427).

REFERÊNCIAS

ANDRADE, Amália. *Um Estudo Experimental das Vogais Anteriores e Recuadas em Português*. 1987. 355 p. Dissertação (Mestrado em Linguística Portuguesa para acesso à categoria de Investigador Auxiliar) - Centro de Linguística da Universidade de Lisboa, Lisboa.

BRYMAN, Alan; CRAMER, Duncan. *Quantitative Data Analysis with SPSS Release 10 for Windows – A Guide for Social Scientists*. Hove: Routledge, 2001. 295 p.

CHILDERS, Donald. *Speech Processing and Synthesis Toolboxes*. New York: John Wiley & Sons, Inc, 2000. 483 p.

de SOUSA, Elizabeth M. G. *Para a Caracterização Fonético-Acústica da Nasalidade no Português do Brasil*. 1984. 170p. Dissertação (Mestrado) - Universidade Estadual de Campinas, Campinas.

DELGADO-MARTINS, Maria Raquel. *Fonética do Português. Trinta anos de investigação*. Lisboa: Edit. Caminho, 2002.

LALWANI, Ajit. *Flexible Formant Synthesizer: A Tool for Improving Speech Production Quality*. 1991. 318 p. Dissertação (Doctor of Philosophy) - University of Florida, Gainesville.

MOUTINHO, Lurdes Castro et al. Contributo para o estudo da variação contextual e regional das vogais nasais do Português Europeu. In: DUARTE, Isabel Margarida; BARBOSA, Joaquim; MATOS, Sérgio; HÜSGEN, Thomas (Org.). *Encontro Comemorativo dos 25 anos do Centro de Linguística da Universidade do Porto*. Porto: Universidade do Porto, 2001 (Julho 2002). 2 v. v. 2, p. 5-17.

TEIXEIRA, António; VAZ, Francisco. European Portuguese nasal vowels: An EMMA Study. In: *Proceedings Eurospeech*. Aalborg, Dinamarca, 2001.

TEIXEIRA, António. *Síntese articulatória das vogais nasais do Português Europeu*. julho 2000. 235 p. Tese (Doutoramento em Engenharia Electrotécnica) - Universidade de Aveiro, Departamento de Electrónica e Telecomunicações.

Alguns esquemas entoacionais característicos da leitura de um texto por portugueses e brasileiros

António Romano

Università di Torino - Itália

Lurdes de Castro Moutinho

Universidade de Aveiro – Portugal

Résumé

Cette étude a pour but immédiat celui de mettre en évidence quelques éléments intonatifs particulièrement typés dans des productions linguistiques communes du Portugais parlé au Brésil (PB) par rapport aux schémas réalisés le plus communément – dans les mêmes conditions – par des locuteurs du Portugais européen (PE).

Le manque d'études dans ce domaine a été une des motivations de cette recherche.

Notre approche privilégie la recherche de configurations mélodiques associées à des figures expressives manifestées sur l'échelle des

contours d'unité tonale et intonative avec un aperçu qui tienne en compte des variables de structuration plus globales.

Notre étude est basée sur un texte lu-interprété par 6 locuteurs (3 locutrices portugaises de Aveiro et 3 locuteurs brésiliens).

Nos résultats montrent qu'en PB émerge assez systématiquement profil final de modalité déclarative, présentant des caractéristiques qui peuvent donner lieu à une sorte de cliché mélodique. Ce contour, ancré à des rapports de hauteur très systématique, se manifeste par une tonie caractérisée par une brusque montée mélodique sur la préaccentuelle, suivie d'une diminution abrupte de hauteur qui se stabilise ensuite sur la voyelle accentuée, pour se réduire encore sur les voyelles postaccentuelles. Cette tonie est précédée souvent par une longue séquence quasi-monotone qui réalise une protonie plate assez aiguë, généralement inconnue au PE.

1. INTRODUÇÃO

Esta comunicação tem como principal objetivo evidenciar os elementos entoacionais, particularmente tipificados em produções linguísticas comuns, do português falado no Brasil (PB) e relacioná-los com enunciados produzidos por falantes do português europeu (PE). A escassez de estudos neste domínio foi a principal motivação para esta pesquisa.

Com efeito, para estas duas variedades do Português, para além da caracterização das particularidades no nível segmental, têm constituído objeto de estudo privilegiado, as diferenças evidentes ao nível da estrutura rítmica e da organização temporal (BARBOSA, 2000; FROTA & VIGÁRIO, 2000 e 2001; GALVES et al, 2002). No entanto, e apesar de pesquisas já efetuadas neste domínio específico (MIRA MATEUS, 1990; MORAIS BARBOSA, 1994), uma comparação eminentemente entoacional entre o PE e o PB revela-se ainda incipiente.

Assim, e numa perspectiva de estudo da variação geolingüística, enquadrada num projeto mais vasto da análise prosódica das variedades românicas (ROMANO, 2001a, CONTINI et al. 2002), pretendemos testar macroíndices de diferenciação prosódica entre PE e PB.

Enquanto a maior parte dos estudos recentes sobre entoação se consagram, inexplicavelmente, à detecção de tipos acentuais, baseando-se unicamente em parâmetros de altura melódica (mesmo para as variedades em que os acentos quantitativos parecem também importantes), e em coerência com o nosso objetivo de estudar a entoação, a nossa perspectiva de análise privilegia a detecção e análise de configurações melódicas associadas a representações gráficas que se apresentam como significativas e que se manifestam através de uma escala de contornos de unidades tonal e entoacional,

sem perder de vista outras variáveis que dão conta de uma estruturação mais global (cf. 't HART et al., 1990; ROMANO, 2001b).

A perspectiva de interpretação dos movimentos melódicos observados visa reconhecer uma estruturação dependente, não só de um certo número de restrições sintáticas, semânticas, pragmáticas e lógicas, mas também de parâmetros de estruturação e distribuição da informação, que podem fazer variar o contorno daí resultante, deixando, no entanto, a possibilidade de ter em conta contornos expressivos que o falante escolhe (ou não) manifestar.

Se, por um lado, esta perspectiva de análise segue as teorias mais confiáveis sobre a estruturação prosódica e sobre a relação entoação-sintaxe (MARTIN, 1977, ROSSI, et al. 1981, AUBERGÉ, 1991, HIRST & DI CRISTO, 1998, ROSSI, 1997), por outro lado, não podemos, no entanto, negligenciar a existência de restrições puramente prosódicas, de restrições eufônicas e eurítmicas (cf. também MARTIN, 1997) e, evidentemente outros de natureza emocional (FONAGY, 1983, CAELEN-HAUMONT, 2002).

Com efeito, cada vez que uma mesma frase é produzida, entram em jogo todas as variáveis, com o objetivo de respeitar “os encontros estruturais”, tal como são definidos por Aubergé (AUBERGÉ, 1991), embora as mesmas variáveis sejam também manipuladas inconscientemente pelo falante que introduz informações extra ou paralingüísticas. Dentre elas, indícios mais ou menos evidentes e mais ou menos controláveis, da sua proveniência geo-sociocultural.

É nesta linha que se enquadra a nossa análise, cujo primeiro objetivo é testar o espaço de variabilidade de dois macro-subsistemas prosódicos do português (PE e PB) e de observar/descrever toda a variabilidade interna que estes agrupamentos podem apresentar.

Com esse fim, propomo-nos começar por recolher “melodias de fala” exploradas enquanto elementos de caracterização e identidade linguística do falante e, por isso, particularmente freqüentes nas produções dos indivíduos de diferentes comunidades. O objetivo é seguramente ambicioso, tanto mais que se inscreve no quadro de uma investigação tipológica que poderia exigir longas

operações de reconhecimento prévio nas diferentes áreas lingüísticas, podendo os isófonos prosódicos não coincidirem com as tradicionais fronteiras dialetais.

No nosso caso, este pré-inquérito, apenas iniciado, permitiu-nos evidenciar somente um traço de diferenciação bastante estável. A identificação desse traço diferenciador é sempre possível – na ausência de um julgamento prévio, por parte de dialetófonos da área lingüística em questão –, se procedermos a um treino perceptivo de sensibilização para este tipo de fenômeno e na condição de encontrarmos uma amostragem de produções que contenha, com alguma clareza, esse tipo de traço.

O trabalho aqui apresentado tem por base um *corpus* reduzido, quase um estudo de caso, obtido a partir de um texto lido, de forma expressiva, por seis falantes: três portugueses e três brasileiros, devendo a sua análise permitir-nos encontrar algumas características de diferenciação entre o PE e o PB. Uma análise preliminar do referido *corpus* deixa já antever uma preferência por esquemas melódicos específicos, claramente distintos, apresentando-se mais ou menos estáveis, para as duas variedades, nas mesmas posições de produção e com estruturas segmental, sintática e logico-pragmática similares.

2. ENQUADRAMENTO TEÓRICO E METODOLOGIA

Trabalhos específicos têm norteado a história de investigações prosódicas sobre o Português e que, ao serem perspectivados de forma distinta – descritiva, contrastiva ou de representação fonológica –, permitiram aprofundar diversos aspectos neste domínio, utilizando metodologias diversificadas (para o PE v. DELGADO MARTINS, 1982; para o PB v. RAMEH, 1962 e 1966 e, no âmbito de estudos mais gerais, mas marcados pela escolha clara de uma teoria, LOPEZ STRODT, 1979 et MAJOR COLEMAN, 1979).¹

Testemunho de um interesse renovado neste domínio é a obra recente de D. Hirst e A. Di Cristo (1998) que, utilizando as mesmas bases e a mesma metodologia de análise, apresentam um estudo para as

diferentes línguas e seus espaços linguísticos respectivos e onde podemos encontrar duas seções especialmente consagradas à apresentação dos sistemas prosódicos do PE (da autoria de M. Cruz Ferreira) e do PB (da autoria de J.A. de Moraes). No entanto, e dado o objetivo da obra, poucas informações fonéticas foram previstas e, por isso, poucos elementos de comparação, neste domínio, estão disponíveis. Em segundo lugar, à parte o espaço utilizado para a apresentação de precisões terminológicas e uma escolha diferente relativamente aos aspectos aprofundados, também a diferente estrutura adotada pelos dois autores nessas seções consagradas ao português, dificultam a recuperação de informações essenciais quando o objetivo é o de uma análise entoacional comparada e numa perspectiva variacionista. Apesar disso, observações incontornáveis estão presentes na contribuição de De Moraes (cf. outras publicações do autor, cujas conclusões se encontram resumidas nesta referência), relativamente ao sistema prosódico do PB, sempre sustentadas por dados quantitativos.²

Utilizando um método rigoroso que aplicou ao estudo e comparação dos sistemas fonéticos das diferentes línguas e variedades, o único investigador, tanto quanto nós sabemos, que tentou, até o momento, apresentar índices de diferenciação para o PE e o PB, parece ter sido o foneticista italiano L. Canepari (cf. CANEPARI, 1983 e 1985). Apoiando-se numa análise auditiva, a maior parte do tempo baseada sobre a sua própria percepção de um certo número de características, generalizadas a partir de um material diversificado, este autor soube resumir (no âmbito de publicações gerais), em alguns traços, as principais constantes que se podem esquematizar, quando de uma comparação ideal entre PE e PB. No seu percurso, seguindo uma tradição de análise entoacional de escola britânica, apresenta – sobre uma espécie de escala musical chamada *tonograma* – esquematizações de perfis melódicos tal como são percebidos pelo ouvido, com indicação de agrupamentos temporais, onde as proeminências são realizadas com uma certa estabilidade e, no caso das línguas com acento livre – acento lexical –, a indicação das posições dos segmentos acentuados. Geralmente, onde, tal como em

outros modelos atualmente existentes, podemos distinguir posições *nucleares* e *pré-nucleares*, esta análise visa a descrição do desenvolvimento de uma fase preparatória, denominada *protonia*, encaminhando-se para uma fase terminal, *tonia*, marcada habitualmente por um perfil específico, como o apresentado nas figuras 1 e 2.³

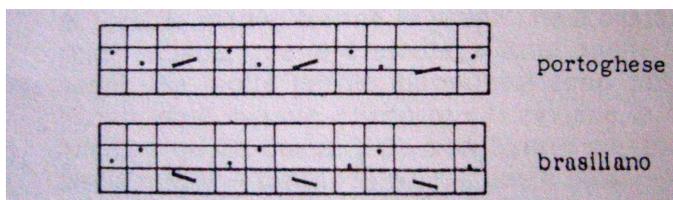


Fig. 1. Esquemas característicos de realizações da protonia em PE e PB, tal como foram descritos por L.Canepari (1985, p. 41).

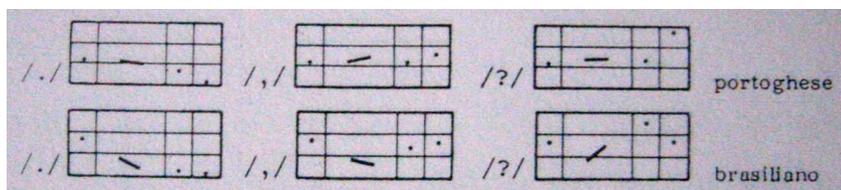


Fig. 2. Esquemas característicos de realizações de diferentes tonias em PE e PB, tal como foram descritos por L.Canepari (1985, p. 50).

Tendo em conta o que nós pudemos observar, esta divisão merece que se dê mais atenção ao desenvolvimento melódico na transição entre as duas partes – a *tonia* e a *protonia* – que constituem a *intonia* que revela, por si só, propriedades específicas. Um outro aspecto digno de maior atenção, negligenciado por uma escolha prévia de quantificação por níveis, diz respeito às relações melódicas a ser estudada em termos de configurações melódicas, ligadas a níveis freqüenciais. Estas podem determinar contornos musicais muito tipificados, mas que se tornarão imperceptíveis numa modelização por níveis ou por finais binários (cf. 't HART et al. 1990).

Por outro lado, um dos problemas que permanece é a generalização de fatos observados, com base em material variável, tendo-se, de início, efetuado uma escolha de fatores considerados responsáveis

pela caracterização prosódica. Este tipo de abstração, pelo menos no nosso caso, conduz ao aparecimento de contornos estilizados, para os quais será necessário confrontar as duas dimensões de alinhamento e a amplitude.

Além disso, e ao invés do que acontece com as representações binaristas para a descrição de fenômenos prosódicos (focalizado/não focalizado, alto/baixo, etc.), quando as variáveis parecem, pelo contrário, ser do tipo linear ou pelo menos não-binárias (cf. CAELEN-HAUMONT, 2002, ROMANO, 2001b), tomamos outro procedimento. Decidimos não negligenciar o fato de que os contornos detectados e estudados, em função de uma diversa interpretação expressiva manifestada em diferentes momentos pelos falantes gravados, sejam o resultado de uma especificidade geoprosódica de um certo número de melodias específicas e de estratégias particulares de estruturação entoacional da informação (exemplos disponíveis para dialetos do PE, nos primeiros resultados do projecto AMPER-Portugal, v. MOUTINHO et al., 1999; para a comparação PE/PB, v. FROTA, 2002).

Apesar disso, a divisão *protonia/tonia*, que aplicamos a uma análise instrumental, revelou-se muito útil na comparação das diferentes realizações (inclusive o suporte constituído pelas análises já publicadas por Canepari), mas o material de que dispomos apenas nos permitiu confirmar uma parte desses esquemas.

As realizações por nós analisadas possibilitaram a observação de elementos importantes de caracterização que, dificilmente, poderiam manifestar-se em gravações laboratoriais, efetuadas com falantes profissionais, nos quais estes, certamente, reproduziriam a variedade de prestígio, controlando-a também no nível expressivo.

O nosso *corpus* reduzido, utilizado para este estudo comparativo, foi gravado a partir de um texto lido-interpretado por 6 falantes (3 portuguesas de Aveiro e 3 brasileiros, um do Rio de Janeiro, outro da Bahia e uma de São Paulo, atualmente residente em Aveiro).⁴ O texto gravado inspira-se num excerto de um curso de Português para aprendizes franceses (DIAS DA SILVA J. & PARVAUX S., 1991), sendo constituído por nove unidades tonais.

1. Meu amigo, os cruzados já acabaram.
2. Então você não sabe que até o nome da moeda varia no Brasil?
3. Antes de 1986 era o cruzeiro,
4. Depois o cruzado,
5. Em 1988 o cruzado novo...
6. E voltou a chamar-se cruzeiro em 1990.
7. Eles só trazem dólares americanos.
8. Nenhum brasileiro viaja com cruzeiros.
9. Você sabe qual é a cotação do dólar?

Para cada uma das realizações (1 a 3 para as 9 unidades indicadas e para cada um dos falantes dos dois grupos), seguimos as etapas de análise e de modelização utilizadas no âmbito do projeto AMPER, recorrendo especialmente a *Speech Analyzer e Matlab*.⁵ Assim, a análise previu:

- extração dos valores de F_0 ;
- estudo da variação de F_0 ao longo da produção de cada unidade tonal, subdividida em protonia e tonia;
- observação da influência do acento lexical, no interior de cada unidade tonal, por considerarmos ser um índice relevante no caso de línguas de acento móvel;
- detecção e modelização de melodias específicas das produções analisadas;
- comparação/interpretação dos resultados obtidos para cada uma das variedades em estudo.

Relativamente a este último ponto, que constitui o objetivo central desta pesquisa, ainda não podemos fornecer resultados, mesmo que parciais, a não ser os relativos a macroíndices que separam os dois supersistemas do PE e do PB.

3. INTERPRETAÇÃO DOS RESULTADOS

Após uma observação atenta das curvas melódicas, as tendências intragrupo aparecem generalizadas, enquanto que as causas de diferenciação, entre as produções dos falantes do mesmo grupo, foram excluídas desta descrição. Através das realizações dos diferentes falantes, constatamos que é a primeira parte da unidade 1 (aqui indicada como 1^a), com um esquema melódico do tipo vocativo (“meu amigo”, facilmente transformável numa frase do tipo “oh meu amigo!” – vocativo, tomada de palavra, ironia, etc.) que apresenta uma maior variabilidade melódica. Os contornos das interrogativas (2. e 9.) revelaram-se particularmente decepcionantes – normalmente muito interessantes no plano de comparações geoprosódicas –, porque foram provavelmente interpretadas de forma bastante diferente, por cada um dos falantes. Mais estáveis, mas igualmente não muito marcadas no nível de uma diferenciação no plano regional dos enunciados, apresenta-se a evolução das curvas melódicas expressas em conformidade com as unidades 3., 4. e 5. Soluções bastante coerentes foram escolhidas, por parte dos falantes portugueses, para a realização dos esquemas declarativos (1b., 7., 8.), tendo-se revelado mesmo muito interessante a interpretação do enunciado 6.

Inscrita no quadro de uma progressão de mudança de movimentos, esta frase podia ser interpretada como uma forma excessiva, quase limite, no nível da sua produção: da mesma maneira, o falante do Rio, influenciado pela evidência referencial do conteúdo, carrega-a de uma certa ironia, realizando um esquema facilmente identificável. Este tipo de realização está também presente na fala da entrevistada de S.Paulo, embora de forma um pouco atenuada.

Este perfil de modalidade, que emerge com bastante sistematicidade em certas variedades do PB, manifesta-se por uma brusca subida melódica, sobre a pré-acentuada, seguida de uma diminuição abrupta de altura, estabilizando-se em seguida sobre a vogal acentuada, para se reduzir ainda nas vogais pós-acentuadas (encontramos aqui o contorno

estilizado proposto por CANEPARI, 1983, v. fig. 2). Esta realização da *tonia* é precedida, freqüentemente, por uma longa seqüência quase monótona e que realiza um contorno final de *protonia* bastante agudo (v. fig. 3 e traçado em negrito na fig. 4).⁶ Este contorno, convertido num estilema prosódico típico do PB,⁷ sofreu diversas manipulações, para testarmos os aspectos que melhor o caracterizavam.

Após experiências informais de simulação por síntese de contornos diferentes, gerados a partir deste, uma certa rigidez específica foi observada: uma diminuição ou aumento de F_0 , superior a 5Hz, pode ser responsável pela perda da força deste contorno. Com efeito, observamos que, nas realizações menos expressivas, o desvio entre os elementos, caracterizados por uma certa relação com a altura média, diminui e que certos valores podem não ser atingidos, em diferentes momentos onde seriam esperados, tomando por base a realização mais expressiva de que dispomos.

A configuração mais caracterizadora do PB que encontramos apresenta-se, por isso, plana (correspondente à seqüência monótona realizada, *steady stretch*), cerca de 4 qT (4 quartos de tom) acima da frequência média do falante, seguida de um pico (pré-acentual) a +8 qT (sempre relativamente à altura de referência), seguidamente assistimos a uma descida bastante abrupta de -8 qT, incidindo sobre a vogal acentuada (com maior duração), e uma descida gradual sobre as vogais pós-acentuadas, atingindo o valor máximo de -15 qT.

Numa simulação com uma melodia sintética, obtida a partir dos valores atingidos para o falante do Rio – o mais facilmente identificado – pareceu-nos que esta configuração passa por quatro níveis estáticos de referência, respectivamente a 180 Hz, 208 Hz, 128 Hz et 102 Hz, com uma oscilação entre os 3 a 4 Hz. Evidentemente que, no que diz respeito ao contorno da frase espontânea, ele apresenta uma evolução dinâmica, com um alinhamento bem específico dos aspectos melódicos mais destacados com a estruturação segmental da frase.

De seguida, apresentamos, na fig. 3, o contorno melódico tal como foi medido para realizações em PB, com correspondências segmentais em transcrição fonética; na fig. 4, alguns contornos sobrepostos,

extraídos de produções portuguesas e brasileiras, relativas a uma parte da protonia e da *tonia* da mesma frase e que mostram, também, uma certa intervariabilidade; finalmente, a fig. 5 apresenta uma comparação de versões estilizadas (cf. ‘t HART et al. 1990) de curvas melódicas realizadas com os mesmos esquemas entoacionais.

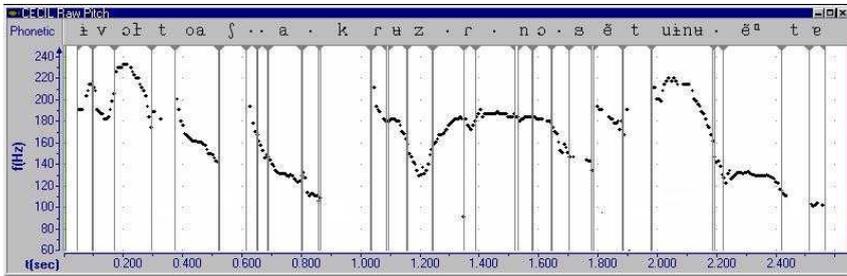


Fig. 3. Curva extraída para a unidade entoacional 6 (“E voltou a chamar-se cruzeiro em 1990”) realizada enfaticamente pelo falante brasileiro do Rio. [gráfico obtido com *Speech Analyzer - SIL*].

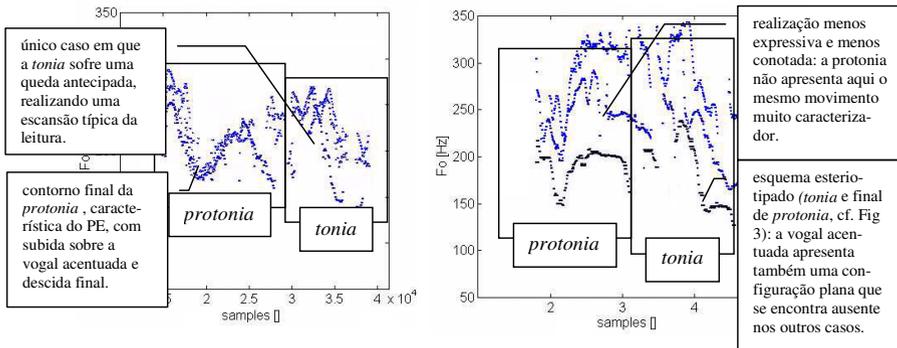


Fig. 4. Contornos sobrepostos de produções portuguesas e brasileiras relativas à mesma passagem da Fig. 3 [gráfico obtido com Matlab – metodologia de extração de F_0 do autor AR].

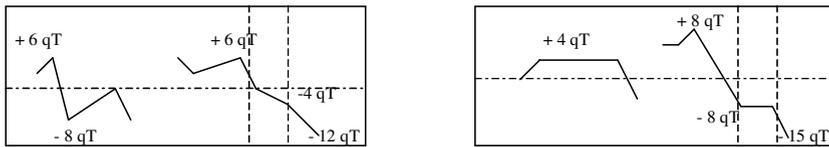


Fig. 5. Contornos estilizados de esquemas mais característicos de PE e PB em correspondência com (a.) a parte final da protonia e (b.) a tonia na interpretação dada por brasileiros e portugueses da mesma passagem da Fig. 3 [esquema feito manualmente]

4. CONCLUSÕES

Numa reflexão prévia sobre as razões que conduziam à diferença entoacional entre PE e PB, constatamos, embora com um número reduzido de falantes, uma certa estabilidade na produção de alguns esquemas melódicos específicos aos dois supersistemas. A ordem da variabilidade por vezes constatada (quando se trata de pessoas do mesmo país, da mesma cidade e do mesmo nível social) apresenta como única explicação as diferenças nas escolhas de soluções expressivas para interpretar e transmitir o mesmo conteúdo. Esta estratégia utilizada por parte dos locutores deve ser, por isso, considerada prioritária nos estudos sobre a entoação; os investigadores que trabalham sobre este aspecto da estruturação linguística da mensagem, deveriam tê-la em consideração no momento da descrição da prosódia de uma qualquer variedade linguística. Certamente que outros fatores de variação dialetológica e sociolinguística deverão ser considerados (cf. MOUTINHO et al., 1999, ROMANO, 2001b), podendo estes funcionar, em alguns casos, como complemento de informação, no momento da leitura dos resultados.

Apesar de não termos perdido de vista essas variáveis, no decorrer desta pesquisa, o estudo contrastivo realizado permitiu-nos apenas destacar um esquema de variação entoacional, de entre o grande número que poderia ser identificado e organizado tipologicamente. Este esquema, do tipo declarativo, que aqui aparece com mais

freqüência nas produções de um falante do Rio de Janeiro, nas marcadas por ênfase expressiva, poderia estar associado a um melisma identificativo, a uma espécie de melodia específica, uma entoação estereotipada, de um PB, em oposição ao PE.

Concluimos, deixando aqui uma observação metodológica relativa às simplificações efetuadas sobre a representação das variáveis prosódicas. Na verdade, com os dados que analisamos, pareceu-nos inútil recorrer a um sistema de transcrição larga ou a um instrumento de etiquetagem prosódica do tipo interpretativo. Enquanto não tivermos ainda testado e inventariado as múltiplas possibilidades de variação, afigura-se-nos prematuro escolher um esquema arbitrário que generalizássemos como modelo de base. Descrever todos os pormenores, fazendo uma escolha meramente acidental, ou ainda mais discutível, introduzir técnicas de simbolização específicas, como se se tratasse de invariantes, conduzir-nos-ia, conseqüentemente, a interpretações errôneas dos resultados, daí a nossa opção.

Queremos também deixar claro que o estudo de caso, aqui apresentado, deve ser entendido como um estudo preliminar num domínio de pesquisa de grande complexidade. Aprofundar esta temática é um desafio que nos propomos assumir num futuro muito próximo.

NOTAS

¹ Para fenômenos específicos, analisados comparativamente entre o PE e o PB, trabalhos recentes abordam o problema de um ponto de vista instrumental e recorrendo a novos modelos (entre outros, FROTA & VIGÁRIO, 2001; FROTA 2002).

² Num capítulo conclusivo, o autor acrescenta ainda considerações importantes sobre a variabilidade geográfica deste sistema, propondo, simultaneamente, um esboço de comparação entre certos traços de variedades geográficas do PB e do PE.

³ Utilizaremos, ao longo de todo o texto, as expressões de Canepari – *Tonia*, *Protonia* –, na acepção que este autor lhes atribui, evitando a tradução por *Tônica*, etc., que não evocam o mesmo conceito. Na verdade, *Tonia* e *Protonia* referem-se a agrupamentos silábicos, ao contrário do que acontece com os conceitos de *Tônica* e *Pretônica* que se referem apenas a vogais ou sílabas isoladas.

⁴ Outros falantes dos dois sexos (3 mulheres e 1 homem), de diferentes variedades do Português europeu (Lisboa, Barlavento, Açores), foram também gravados para validar a representatividade da leitura que foi feita do texto em que se baseia esta análise. Os dados destas produções constituem uma análise em curso, mas uma leitura sumária das curvas melódicas aqui apresentadas confirma as tendências observadas e resumidas neste artigo.

⁵ Com a utilização de aplicações realizadas, no âmbito do referido projeto, por um dos autores deste artigo, AR.

⁶ Este movimento bastante sistemático da *protonia* do PB escapa, no entanto, às descrições dadas por Canepari (1985); cf. Fig. 2.

⁷ Há grandes possibilidades de reencontrar esta configuração melódica em entoações com conteúdo e conotação pragmático-lingüística similar no dialeto italiano de Gênova e nas variedades do italiano regional existente na Ligúria (Itália do Norte).

REFERÊNCIAS BIBLIOGRÁFICAS

AUBERGE, V. La synthèse de la parole: “des règles aux lexiques”. 1991. Tèse (Doutoramento) – ICP-INPG/Université Stendhal et CRISS-Université Pierre Mendès-France, Grenoble. 1 v.

BARBOSA, P.A. Illuminating some Methodological Issues concerning Speech Timing Research from a Comparison between European and Brazilian Portuguese. *Cadernos de Estudos Lingüísticos*, n. 39, p. 41-50, 2000.

CAELEN-HAUMONT, G. Perlocutory values and functions of melisms in spontaneous dialogue. *Proc. of the Int. Conf. Speech Prosody 2002*, Aix-en-Provence, p. 195-198, 2002.

CANEPARI, L. *Phonetic notation / La notazione fonetica*. Venezia: Cafoscarina, 1983. (vol. com dois cassetes).

CANEPARI, L. *L'intonazione. Linguistica e paralinguistica*. Napoli: Liguori, 1985.

CONTINI M.; LAI, J.P.; ROMANO, A.; ROULLET, S.; MOUTINHO, L. de C.; COIMBRA, R.L.; PEREIRA, B. U.; SECCA, Ruivo S. Un Projet d'Atlas Multimédia Prosodique de l'Espace Roman. *Proc. of the Int. Conf. Speech Prosody 2002*, Aix-en-Provence, p. 227-230, 2002.

CRESTI E. Information and Intonational Patterning in Italian. In: FERGUSON, B. et al. (Ed.). *Accent, Intonation et modèles phonologiques*. Toronto: Mélodie, 1994. p. 99-140.

- CRUZ-FERREIRA, M. Intonation in European Portuguese. In: HIRST, D.; DI CRISTO, A. (Ed.). *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press, 1998. p. 167-178.
- DE MORAES, J.A. Intonation in Brazilian Portuguese. In: HIRST, D.; DI CRISTO, A. (Ed.). *Intonation Systems. A Survey of Twenty Languages*. Cambridge: Cambridge University Press, 1998. p. 179-194.
- DELGADO MARTINS, M.R. (1982). *Sept Études sur la Perception. Accent et Intonation du Portugais*. Lisboa: INIC, 1986.
- DIAS da SILVA, J.; PARVAUX, S. *Portugais. Pratiquer. Perfectionnement: pour maîtriser la langue et son environnement*. Les langues pour tous, Presses Pocket, 1991. (vol. com très cassetes.)
- FONAGY, I. *La vive voix. Essais de Psychophonétique*. Paris: Payot, 1983.
- FONAGY, I.; FÓNAGY, J. L'intonation et l'organisation du discours. *Bull. de la Société de Linguistique de Paris*, LXXVIII, 1, p. 161-209, 1983.
- FONAGY, I.; L.J., Boë; LUCCI, V. Incidence des variables situationnelles sur quelques paramètres de l'intonation. *Actes des VIII^{èmes} JEP-GALF*. Aix-en-Provence, 1977, p. 13-19.
- FROTA, S. The prosody of focus: a case-study with cross-linguistic implications. *Proc. of the Int. Conf. Speech Prosody 2002*. Aix-en-Provence, 2002. p. 319-322.
- FROTA, S.; VIGÁRIO, M. Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. *Actas do XV encontro da APL*. Braga, 2000. p. 533-555.
- FROTA, S.; VIGÁRIO, M. On the correlates of rhythm distinctions: the European/Brazilian Portuguese case. *Probus*, n. 13, p. 247-275, 2001.
- GALVES, A.; GARCIA, J.; DUARTE, D.; GALVES, C. Sonority as a basis for rhythmic class discrimination. *Proc. of the Int. Conf. Speech Prosody 2002*. Aix-en-Provence, 2002, p. 323-326.
- ‘T HART J.; COLLIER, R.; COHEN, A. *A perceptual study of intonation*. Cambridge, Cambridge University Press, 1990.
- LOPEZ STRODT, B. The Sound Pattern of Brazilian Portuguese (Cariocan Dialect). 1979. Dissertation (Ph.D.) – University of California, Los Angeles.
- MARTIN, Ph. Syntax and intonation: an integrated theory. *Toronto Semiotic Circle, Monographs Working Papers and Prepublications*, Victoria Univ., Toronto, 1977.
- MARTIN, Ph. La Musique de la Phrase. In: GEZUNDHAJT, H.; MARTIN, Ph. (Ed.). *Promenades en Phonétique*, Toronto, Mélodie, p. 97-112, 1997.

- MAJOR COLEMAN, R. Prosody in Brazilian Portuguese phonology. 1979. Dissertation (Ph.D.) – The Ohio State University.
- MIRA MATEUS, M. H. *Fonética, Fonologia e morfologia do português*. Lisboa: Universidade Aberta, 1990. (em col. com A. Andrade, M. Do Céu Viana & A. Villalva).
- MORAIS BARBOSA, J. Entoação e Prosódia. In: HOLTUS, G. et al. (Ed.). *Lexikon der Romanistischen Linguistik*, Band VI, n. 2, p. 243-248, 1994.
- MOUTINHO, L. De C., COIMBRA, R.L., SECCA, R. S.; PEREIRA, B. U. (2001). Atlas Prosódico Multimédia: curvas de uma trajectória. *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*. Coimbra, 1999, Colibri, 2001. p. 387-391.
- RAMEH, Cl. A intonação no português do Brasil. *Estudos Linguísticos*, São Paulo, dez. 1966.
- RAMEH, Cl. Contrastive analysis of English and Portuguese intonation. 1962. Thesis (Master) – Georgetown University.
- ROMANO, A. Un projet d'Atlas Multimédia Prosodique de l'espace roman. *Actes du XXIII^{ème} CILFR*. Salamanca, 2001a, en cours d'impression (éd. F. Sanchez Miret).
- ROMANO, A. Applicabilité des systèmes de transcription et d'analyse de l'intonation aux cas de variabilité dialectale présentés par la situation géoprosodique italienne. *Actes des Journées Prosodie 2001*, Grenoble, 2001b, en cours d'impression (éd. A. Lacheret-Dujour).
- ROSSI, M. Modèles de perception et de production de l'intonation. In: ROSSI, M. et al. (Ed.). *L'intonation: de l'acoustique à la sémantique*. Paris: Klincksieck, 1981.
- ROSSI, M. L'intonation et l'organisation de l'énoncé. *Phonetica*, n. 42, p. 140-144, 1985.
- ROSSI, M. Is syntactic structure prosodically retrievable? *Proc. of Eurospeech '97*, Rhodes, 1997, p. 1-8.

Análise comparada de três padrões prosódicos em francês e em português europeu

Lurdes de Castro Moutinho
Universidade de Aveiro

Jean-Pierre Zerling
Université Marc Bloch à Strasbourg

Résumé

Notre objectif est d'effectuer une comparaison entre trois types de structures prosodiques (déclarative, impérative et interrogative) réalisées dans deux langues romanes de type accentuel différent: le français (accent rythmique) et le portugais européen (accent lexical).

Pour chaque langue, le corpus est constitué d'une phrase lexicalement ambiguë prononcée par un locuteur masculin natif. Les phrases dans les deux langues ont une structure globale très voisine, afin de favoriser leur comparaison. Les paramètres retenus pour l'analyse sont la fréquence fondamentale F_0 , la durée et l'intensité des voyelles. La comparaison porte d'une part sur les différences entre les trois structures prosodiques pour une même langue, puis sur les différences inter-langues pour un même patron prosodique.

Les conclusions montrent, au-delà des particularités individuelles propres aux locuteurs, et malgré la nature accentuelle différente des deux langues, que les stratégies prosodiques utilisées sont en fait très voisines pour la déclarative et l'impérative. En revanche, l'interrogative semble réalisée de manière très différente par les deux locuteurs, tous en présentant néanmoins certains points communs.

OBJETO DE ESTUDO

Esta pesquisa teve a sua origem no momento de uma visita de um dos autores deste artigo ao Centro de Línguas e Culturas da Universidade de Aveiro, onde foi realizado um *workshop* sobre análise prosódica, destinado a estudantes de Fonética.

Ela tem como objetivo:

- por um lado, comparar, para uma língua dada, três tipos de estruturas prosódicas:
 - declarativa
 - interrogativa
 - imperativa
- por outro lado, repetir a análise para várias línguas românicas, neste caso o francês (FR) e o português europeu (PE) e, posteriormente, para o espanhol, o italiano, etc., a fim de serem realizados estudos comparativos. Os resultados destes trabalhos poderão ser comparados com os obtidos no âmbito do Atlas multimídia das línguas românicas (Grenoble, projeto AMPER¹).

PROCEDIMENTO EXPERIMENTAL

1. As diferentes etapas

A pesquisa apresenta-se organizada em diversas etapas: Para cada uma das línguas, teve o seguinte procedimento:

- escolha de uma frase para constituição do *corpus*, obedecendo a certas restrições;
- transcrição fonética da frase;

- gravação por um falante;
- previsão dos acentos lexicais e rítmicos;
- previsão das variações melódicas, segundo a terminologia de Delattre (DELATTRE, 1966);
- observação das variações efetivamente realizadas e concordância com as previsões;
- comparação dos três padrões melódicos;
- análise das durações e da intensidade vocálicas.

Seguidamente, e cruzando os dados obtidos para as diferentes línguas estudadas, fez-se a determinação das particularidades e dos pontos comuns próprios a cada uma das línguas.

2. O corpus

Para a constituição do *corpus* do presente trabalho, foram construídas duas frases cujas estruturas, sintática e fonética, são relativamente próximas:

<i>Pateta toca no café</i>	[patetɛ itəkɛnukaifɛ]
<i>Petitout passe à la télé</i>	[pɛtiutu pasalateilɛ]

Como podemos constatar, tratando-se de duas frases ambíguas, apenas a variação prosódica permitirá, no nível da oralidade, a distinção de sentido em cada uma das modalidades escolhidas: declarativa, imperativa e interrogativa.

No plano segmental, as frases são comparáveis, visto ambas conterem o mesmo número de sílabas fonéticas. As sílabas são abertas, de estrutura CV, e, quase sempre, a consoante que integra a sílaba é uma consoante surda, com o objetivo de facilitar a análise. Note-se que, no caso do francês, esta constituição silábica não favorece o alongamento vocálico.

As diferenças fundamentais entre as frases advêm do fato da sílaba acentuada, no caso dos nomes próprios, não coincidir nas duas línguas e o verbo ser dissilábico em português e monossilábico em

francês. No entanto, e com o objetivo de tornar as duas frases o mais semelhantes possível, escolhemos como último elemento da frase em português uma palavra oxítone acentuada e, por isso, na última sílaba, como sempre acontece com o francês. Para além disso, esta opção permite, simultaneamente, evitar a queda da última vogal na frase portuguesa. Na verdade, estudos anteriores (DELGADO-MARTINS, 1982; MIRA MATEUS et al, 1982; MORAIS BARBOSA, 1994) mostram que as vogais átonas do português europeu, quando em final de sílaba ou de frase, sofrem, freqüentemente, um fenómeno de queda. Estudos por nós realizados (MOUTINHO et al. 2001), comprovam esse tipo de fenómeno, sobretudo em frases do tipo declarativo.

3. Os documentos

Para cada uma das línguas, as frases são produzidas várias vezes pelo mesmo *falante* nativo, do sexo masculino. Para esta análise, foram seleccionadas cinco realizações, para cada tipo de frase, consideradas perceptivamente corretas e consideradas representativas das modalidades em estudo. Dispomos, por isso, de um total de $3 \times 5 = 15$ realizações de cada frase, para cada uma das línguas.

As gravações foram realizadas, respectivamente, nos laboratórios de fonética das Universidades de Aveiro (Portugal) e Estrasburgo (França). O sinal gravado foi convertido no formato “.wav”, de modo a poder ser tratado posteriormente de forma semi-automática.

A *segmentação do sinal*, bem como a elaboração de *gráficos*, foram realizadas em Aveiro, utilizando o programa Matlab e tendo sido adotada uma metodologia cientificamente testada (CONTINI et al. 2002, MOUTINHO et al. 2001).²

Os *parâmetros* analisados para cada realização vocálica são a duração, a intensidade e três valores da frequência fundamental, medidos em três pontos distintos: princípio, meio e fim da vogal. Os valores médios obtidos, para as cinco realizações de uma mesma frase pelo mesmo falante, são calculados em seguida. São esses mesmos valores, inscritos nos quadros fornecidos em anexo, que serão utilizados para a análise.

Os *gráficos* representam, para cada frase, as variações sobrepostas da frequência fundamental, F_0 , para todas as vogais realizadas. A linha média de F_0 é calculada a partir de cinco frases, para cada um dos padrões prosódicos.

Outros *gráficos* representam, para cada frase, as variações médias da frequência fundamental F_0 para cada um dos padrões prosódicos. São estas variações médias da curva melódica que serão, em seguida, comparados, quer para dois padrões entoacionais do mesmo falante, quer ainda para o mesmo padrão nas duas línguas.

Finalmente, *diversos gráficos* permitirão comparar as variações de duração e intensidade das vogais.

4. Cálculo do desvio tonal

Para comparar as variações ou os valores da frequência fundamental, optamos por trabalhar em semitons temperados, partindo da fórmula:

$$F_2 = F_1 (12/2)^x$$

que permite calcular, a partir de uma frequência F_1 , a frequência F_2 situada x semitons mais *altos*, permitindo calcular o valor do *desvio tonal x (em semitons)*:

$$x = (12/\log 2) \cdot \log(F_2/F_1) = 39,86 \cdot \log(F_2/F_1),$$

$$\text{que podemos aproximar de } \mathbf{x = 40 \cdot \log(F_2/F_1)}$$

ANÁLISE

1. Considerações preliminares: comparação dos dois locutores

Em primeiro lugar, parece-nos interessante fazer uma breve análise comparativa dos dois falantes. Globalmente, observam-se dois tipos de diferenças sistemáticas:

- registro;
- tessitura;
- velocidade de produção de fala.

• registro

Denomina-se, geralmente, de *registro* a altura característica da voz de um cantor, no interior do seu alcance vocal: *o registro agudo, alto, médio e grave*.

Embora os nossos dois falantes sejam homens, o seu registro é ligeiramente diferente. O quadro que se segue permite comparar a frequência fundamental média, Fo, calculada para cada um deles a partir de todas as vogais ($8 \times 5 = 40$) e para cada tipo de frase. A última linha do quadro indica o desvio médio entre as duas vozes, convertido em semitons temperados.

Quadro 1 – Frequência fundamental média dos dois locutores para as três frases

Fo médio (Hz / nota)	Declarativa		Imperativa		Interrogativa		?Fo (D/Im-In)
falante FR	154	ré#	153	ré#	181	fa	2,9dt
falante PT	122	si	124	si	146	ré	3 dt
?Fo médio (FR/PT)	4 dt		3,6 dt		3,7 dt		

Note-se que o falante português fala, sistematicamente, com uma frequência fundamental média mais baixa de 3,6 a 4 semitons do que o falante francês, independentemente do tipo de frase.

Por outro lado, a frase interrogativa é produzida, globalmente, pelos dois falantes, com uma voz mais elevada de 3 semitons, relativamente ao que acontece com as frases declarativa e interrogativa (última coluna).

• a tessitura

Devemos assinalar que, no presente trabalho, atribuiremos a *tessitura* um sentido mais amplo do que normalmente lhe é atribuído: o alcance de uma voz que vai da nota mais grave à nota mais aguda e a escala de sons que podem ser emitidos por uma voz, sem qualquer dificuldade.

No quadro seguinte, indicamos, para cada falante e para cada tipo de frase, os valores extremos atingidos para a frequência fundamental (em Hz), o equivalente aproximado em termos de notas

musicais, bem como o desvio correspondente, em semitons, calculado segundo o método anteriormente descrito.

Em seguida, avaliamos o desvio em semitons entre os dois falantes, respectivamente para o Fo mais baixo (mínimo) e o Fo mais alto (máximo), bem como o intervalo entre esses valores.

O gráfico ilustra, com base nos valores inscritos no quadro, em termos de notas musicais, a tessitura própria a cada um dos falantes.

Quadro 2 – Valores extremos da freqüência fundamental e tessitura dos dois falantes para as três frases

		Declarativa	Imperativa	Interrogativa
Falante FR	Fo mini a maxi notas mini a maxi tessitura em dt	94 à 227 Hz Fa1#+ a La2+ = 15,3 dt	101 a 221 Hz Sol1 a La2 = 13,6dt	125 à 270 Hz Si1 a Do3 = 13,3 dt
Falante PT	Fo mini a maxi notas mini a maxi tessitura em dt	91 a 144 Hz Fa1#- a Do2#+ = 7,9 dt	92 a 163 Hz Fa1# a Ré2#+ = 9,9 dt	108 a 185 Hz Sol1#+ a Fa2# = 9,3 dt
Desvios FR/PT	?Fo mini ?Fo maxi ? tessitura	0,6 dt 7,9 dt 7,4 dt	1,6 dt 5,3 dt 3,7 dt	2,5 dt 6,5 dt 6,8 dt

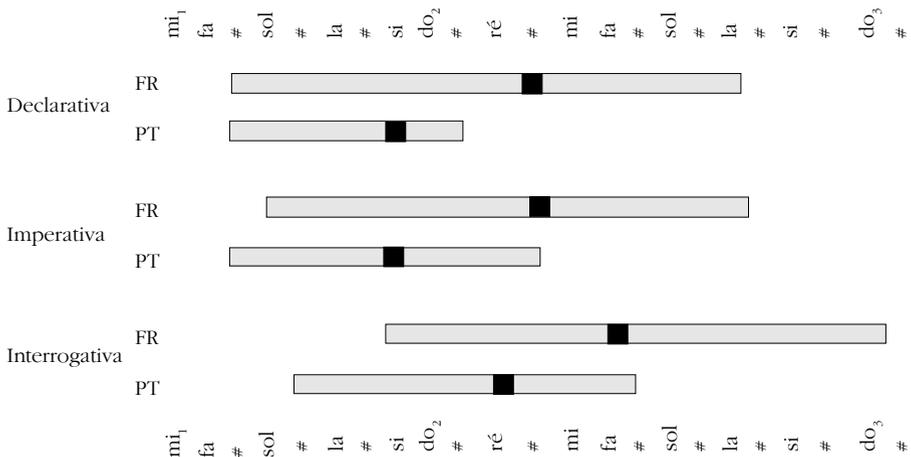


Figura 1 – Tessitura média dos dois falantes para a produção das três frases (A parte mais escura na barra indica a medida média de Fo)

As notas musicais emitidas pelos informantes são bastante próximas (Fa1#) e levar-nos-iam a classificá-los, a ambos, se nos reportarmos aos critérios relacionados com o canto, como barítonos.

Para além disso, o fato mais marcante que pode ser observado no quadro 2 é a diferente tessitura que caracteriza a voz de cada um dos locutores, variando de 3,7 a 7,4 semitons. A tessitura do falante francês, compreendida entre 13,3 e 15,3 semitons, é sistematicamente superior, atingindo o dobro da do português, compreendida entre 7,9 e 9,9 semitons. Este último apresenta-se, por isso, com uma entoação mais monótona.

As considerações que acabamos de tecer devem ser relacionadas com as observações efetuadas a partir dos valores médios de F_0 , inscritos no quadro precedente. Os valores médios superiores, relativamente ao falante francês, não refletem uma voz sistematicamente com colocação mais elevada, visto que os valores inferiores dos dois informantes são próximos entre si, mas refletem sobretudo uma grande variabilidade da frequência fundamental.

• o débito e as durações

A audição e a observação do sinal acústico gravado revelam a presença de uma pausa entre dois grupos rítmicos, somente no caso da frase imperativa, em todas as vezes que ela é realizada por qualquer um dos falantes.

O quadro 3 resume o conjunto dos dados: podemos observar a duração total de cada uma das frases, a duração do tempo de pausa, a relação entre a pausa e o tempo de elocução (duração total), a média destes valores, para cada falante e, finalmente, o cálculo da velocidade de elocução (com as pausas), bem como a duração do sinal de fala (sem pausas).

Constatamos que o falante português articula os sons um pouco mais rapidamente que o falante francês (6,7 sílabas por segundo e 6,4 sílabas por segundo, respectivamente). Por outro lado, a pausa efetuada pelo primeiro é claramente mais longa (17,1% da duração total da frase, para o português, 5,3%, para o francês). No entanto, se tivermos em

conta as pausas, é o falante francês que terá uma velocidade de elocução global ligeiramente mais rápida, com 6 sílabas por segundo, em oposição ao português que realiza 5,5 por segundo.

Quadro 3 – Frases imperativas – duração total das frases e das pausas

FR				PT			
frase	duração	pausa	pausa/dur	frase	duração	pausa	pausa/dur
1	1376	64	4.6%	1	1416	261	18.4%
2	1263	68	5.4%	2	1446	240	16.6%
3	1358	71	5.2%	3	1467	249	17.0%
4	1320	72	5.4%	4	1456	256	17.6%
5	1320	77	5.8%	5	1446	230	15.9%
Tot	1327	70	5.3%	Tot	1446	247	17.1%
Desvio	43.5	4.9		Desvio	19.0	12.5	
sil/seg	6.0			sil/seg	5.5		
dur.fala	1257			dur.fala	1199		
sil/seg	6.4			sil/seg	6.7		

Esta tendência de um débito global ligeiramente mais rápido, por parte do falante francês, confirma-se no quadro seguinte, onde podemos comparar as durações médias para os três tipos de frase. A diferença é, no entanto, mínima, visto que a variação se situa entre 0,3 e 0, 6 sílabas por segundo.

Quadro 4 – Duração média das frases e velocidade de elocução (+ com pausas)

	FR			PT		
	tempo de fala	velocidade de elocução	velocidade de elocução	tempo de fala	velocidade de elocução	velocidade de elocução
Declarativa	1137	7.0		1247	6.4	
Imperativa	1257	6.4		1199	6.7	
+ pausas	1327		6.0	1446		5.5
Interrogativa	1177	6.8		1243	6.4	

Poderemos verificar, nos quadros fornecidos em anexo, que as durações vocálicas são inferiores para o falante francês, o que não deixa de ser surpreendente, visto que a sua velocidade de elocução

é ligeiramente superior. Constatamos, também, que as durações vocálicas mais baixas são, em geral, e para cada uma das frases, de 40ms para o falante francês e de 30ms para o português. Retomaremos o estudo desse fenômeno, de forma mais aprofundada, um pouco mais adiante.

- **a intensidade**

Globalmente, a intensidade média das vogais dos dois locutores apresenta-se bastante estável, oscilando entre 95 e 105 dB. Em português, há uma tendência para descer em final de frase, o mesmo não acontecendo no francês. Também este aspecto será retomado no desenvolvimento da análise.

2. Estudo das variações melódicas (Fo)

2.1. Frase declarativa

- **previsão do padrão melódico**

Para a declarativa, o padrão entoacional esperado é globalmente descendente (padrão de “finalidade”, segundo Delattre, 1966). Mas as frases por nós analisadas, incluindo, *a priori*, dois grupos rítmicos, o primeiro grupo pode ser ascendente. Para além disso, poderemos ainda esperar outras eventuais variações, relacionadas com os acentos lexicais, para o português, pela utilização de um ou vários dos três parâmetros acústicos: Fo, duração ou intensidade.

- **falante francês**

Os resultados vão de encontro às nossas previsões: verifica-se uma subida melódica de 8 semitons (143-227 Hz) no primeiro grupo. Note-se que, apesar desta subida, cada segmento vocálico se apresenta, freqüentemente, descendente. Este aspecto está ligado a um fenômeno micro-prosódico, muito freqüente em francês: o ataque glotal é reforçado pela presença de uma consoante surda que antecede a vogal. O ponto culminante (227 Hz) é atingido na 3ª sílaba (fim do grupo rítmico), constatando-se, em seguida, uma descida regular, quase linear de 12,6 semitons (195-94 Hz). Relativamente às

durações, verifica-se que as duas sílabas, finais de grupo, são ligeiramente mais longas do que as outras, sobretudo a última que chega quase a atingir o dobro da duração.

- **falante português**

Constatamos igualmente uma subida nas três primeiras sílabas, embora proporcionalmente menor do que em francês: 3,6 semitons (118-145 Hz). A elevada frequência sobre a vogal átona final pode parecer surpreendente, mas confirma o papel que ela desempenha no plano entoacional, tal como acontece em francês. As duas sílabas de *toca* mantêm-se elevadas e encontram-se ao mesmo nível. Assistimos, em seguida, a uma descida progressiva de 7,7 semitons (142-91 Hz) nas três últimas sílabas. Note-se que as três vogais acentuadas são as mais longas, sobretudo a última, sendo a sua duração 2 ou 3 vezes superior à duração observada para as breves. No entanto, e contrariamente a uma expectativa legítima, a frequência fundamental não permite identificar verdadeiramente o que normalmente são considerados “acentos tônicos”.

- **comparação Português/Francês**

A sobreposição da representação das curvas melódicas obtidas mostra claramente a semelhança entre dois padrões entoacionais escolhidos em ambas as línguas.³ Mas notamos igualmente, independentemente da diferença existente, no nível da tessitura, entre os dois informantes, que a subida do primeiro grupo é claramente superior em francês. Este tipo de realização é perfeitamente compreensível, visto que a terceira sílaba, em francês, é, simultaneamente, portadora do acento rítmico e do padrão entoacional de continuação do movimento. Em português, pelo contrário, a presença do acento sobre a segunda sílaba dificulta, certamente, uma subida sobre a terceira, o que poderia provocar uma eventual ambigüidade.

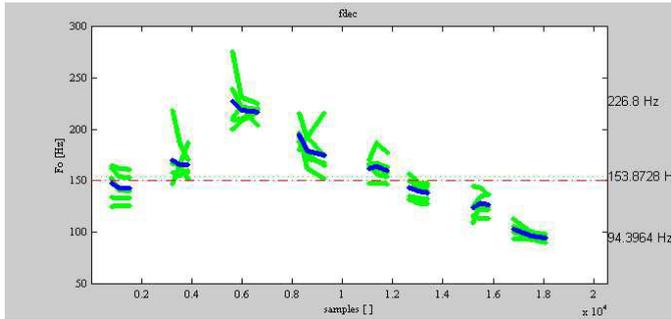


Figura 1 – Média das 5 repetições da frase: Petitout passe à la télé
– francês – Declarativa

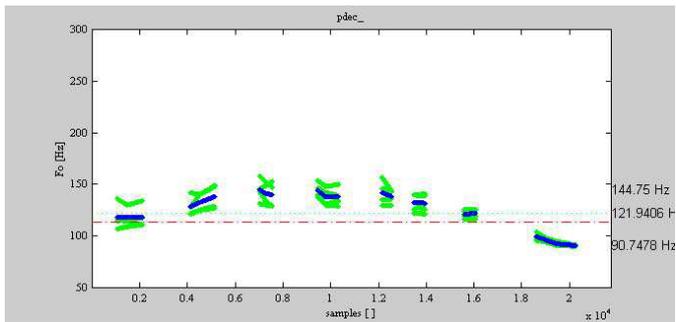


Figura 2 – Média das 5 repetições da frase: Pateta toca no café
– português - Declarativa

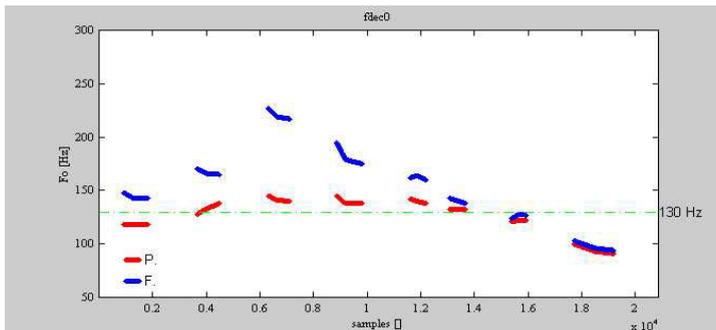


Figura 3 – Comparação das Declarativas PORT/FR

2.2. Frase imperativa

• previsão do padrão melódico

A frase imperativa (a ordem ou “commandement”, segundo a terminologia de Delattre) obedece geralmente a um padrão descendente. No entanto, as nossas frases comportam dois grupos rítmicos, materializando dois sintagmas que, gramaticalmente, correspondem a um vocativo, seguidas de um imperativo. O primeiro elemento (vocativo) poderia perfeitamente ser realizado, tanto por um movimento elevado ou ascendente (do tipo interpelativo), como por um movimento descendente (do tipo exclamativo).

• falante francês

De certa forma interessante, relativamente às nossas previsões, constatamos que a curva melódica do segmento vocativo é globalmente ascendente, 4,3 semitons (130-167 Hz), mas se a terceira vogal é marcada por um ataque elevado, o esquema inverte-se rapidamente e a frequência desce claramente de 2,1 semitons (167-148), o que se torna possível devido a uma longa duração da vogal. Tudo se processa como se tivéssemos, simultaneamente, neste primeiro grupo, uma subida e uma descida combinadas de Fo. A longa duração da terceira vogal reflete, também, sem sombra de dúvida, o fim de um grupo rítmico mais marcado do que na declarativa, tanto mais que esse mesmo grupo é seguido de uma pausa silenciosa.

Podemos admitir que esta descida desempenha uma função contrastiva, visando ainda evidenciar o ataque muito elevado, no que diz respeito à frequência sobre a primeira sílaba do grupo imperativo. A subida brusca de 148 para 221 Hz equivale a 6,9 semitons, o que é verdadeiramente considerável.

A variação melódica do segundo grupo faz-nos pensar na frase declarativa, isto é, com uma descida progressiva de 13,6 semitons (221-101 Hz).

- **falante português**

Podemos dizer que, para este falante, o grupo entoacional que corresponde ao vocativo – 1º grupo rítmico – é, simultaneamente, semelhante e diferente do observado para o informante francês: de fato, constatamos duas variações contraditórias. Se, em francês, a curva melódica é globalmente ascendente com uma descida final, para o português, poderemos dizer que há globalmente uma descida, com uma subida intermédia, necessária para marcar o acento tônico sobre o paroxítono *Pateta*. Esta descida é importante, visto que representa uma diferença de 4 semitons (120-95 Hz) entre o início e o fim deste grupo rítmico, chegando mesmo a atingir uma diferença de 7,1 semitons (143-95 Hz), se tomarmos como ponto de partida, o ponto mais elevado da 2ª vogal. Esta descida é tanto mais importante, por quanto ela conduz a curva melódica ao seu ponto mínimo (95 Hz), que apenas voltaremos a encontrar no final de frase.

O grupo seguinte, como em francês, surge depois de uma pausa (mais longa para este informante) e um ataque ainda mais marcado do que se verifica no falante francês: 9,3 semitons mais elevado (95-163 Hz). Este fato é tanto mais significativo, porque o falante português utiliza menos as variações melódicas do que o francês. Recorde-se que a sua tessitura se situa entre 7,9 e 9,9 semitons. Constatamos que mobiliza, nesta situação precisa, todas as suas capacidades vocais, visto que a sua frequência fundamental passa, praticamente, do seu valor mais baixo ao seu valor mais elevado.

À semelhança do que acontece com a declarativa, verificamos que as três vogais acentuadas são as mais longas, embora, neste caso, seja a segunda, em *Pateta*, que é a mais longa (cerca de 4 vezes mais longa do que as breves). Tal fato é compreensível, na medida em que o vocativo se apresenta como elemento importante. Contrariamente ao que se passa com a declarativa, a frequência fundamental desempenha, neste caso, um papel mais relevante, por permitir destacar os acentos “tônicos”, duas das vezes, através de pontos culminantes e, uma outra vez, pela descida final.

• comparação português/francês

A sobreposição das duas curvas melódicas mostra-se, claramente, e relativizando o efeito tessitura, diferente em cada um dos informantes, à semelhança dos dois padrões entoacionais nas duas línguas. O primeiro grupo apresenta um padrão convexo, com o seu valor máximo sobre a vogal acentuada, variando em função da língua. Em seguida, depois de uma pausa, a frequência fundamental ataca o segundo grupo, com um valor muito elevado nos dois casos. O valor mais elevado em francês é, sem dúvida, condicionado pela já elevada frequência da última vogal do grupo precedente. Recorde-se, no entanto, que, em termos de altura musical, o início deste movimento ascendente é superior em português. O final da frase é, em seguida, progressivamente descendente, constatando-se quase uma sobreposição do traçado para os dois falantes.

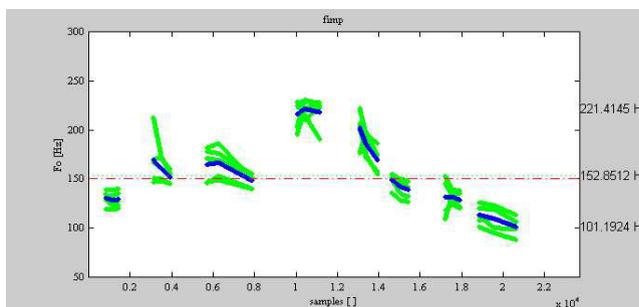


Figura 4 – Média das 5 repetições da frase: *Petitout passe à la télé*
– francês – Imperativa

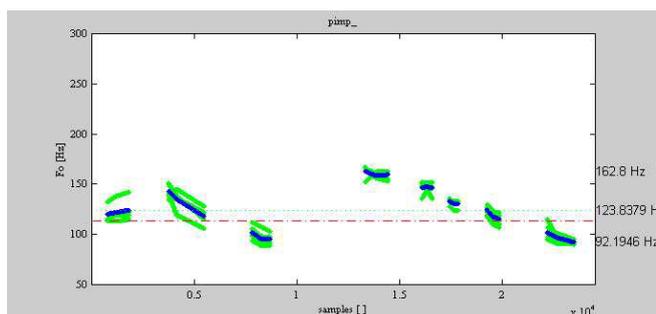


Figura 5 – Média das 5 repetições da frase: *Pateta toca no café*
– português – Imperativa

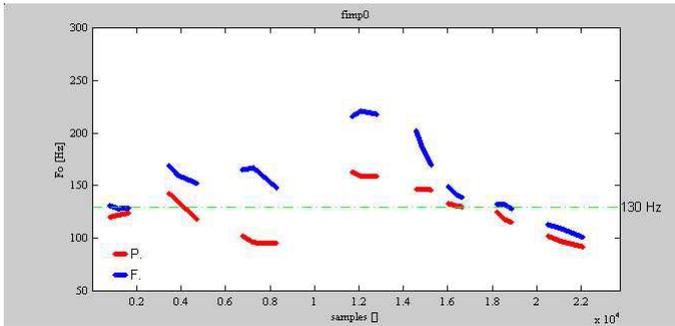


Figura 6 – Comparação das imperativas PORT/FR

2.3. Frase interrogativa

• previsão do padrão melódico

Nas frases por nós analisadas, sendo sintaticamente ambíguas, sabemos que a entoação desempenha um papel necessariamente determinante. Neste caso, estamos perante uma interrogativa global (uma pergunta, “question”, segundo Delattre), o que permite prever uma subida melódica, habitualmente próxima do final de frase. Esta frase sendo, à priori, constituída por dois grupos rítmicos, o primeiro é menos previsível: tanto pode ser ascendente, como descendente, de acordo com a lógica adotada (esquema ascendente de continuação ou descendente, para ser obtido um efeito contrastivo). Acrescente-se a tudo isto as variações acentuais, já referidas.

• falante francês

Como para os dois tipos de frase precedentes, constatamos que a frequência fundamental sobe gradualmente sobre as três primeiras sílabas que constituem a palavra *Petitout*, com uma tessitura de 3,5 semitons (125-153 Hz), próximo da frase imperativa, mas com uma mudança de direção menos acentuada e uma duração mais longa da última vogal do grupo. Em seguida, e depois de uma variação quase nula, na primeira metade do segundo grupo, assistimos a uma subida rápida e abrupta de F_0 sobre a vogal final que se apresenta com uma

longa duração: subida de 8,6 semitons (165-271 Hz), dos quais 5,3 (200-271 Hz) sobre a única vogal! A frequência atingida é a mais elevada das três realizações (271 Hz), resultado que vai de encontro ao nível final proposto por Delattre, para o esquema entoacional de uma pergunta.

• **falante português**

O ataque da frequência fundamental é o mais elevado das três frases produzidas por este falante (161 Hz, em oposição a 118 et 120 Hz). Apresenta-se mesmo como praticamente mais elevado do que os máximos das duas frases precedentes (145 et 163 Hz). Em seguida, depois de uma ligeira subida de 2,4 semitons (161-185 Hz) que culmina na vogal tônica, observa-se uma descida progressiva, muito lenta, até a vogal final que sofre uma queda abrupta de 4,6 semitons (147-113 Hz). No entanto, sensivelmente a meio da sua duração, assistimos a uma inversão do seu movimento, para subir de 1,4 semitons (108-117 Hz). Poderemos admitir que esta subida ligeira, ligada ao movimento de inflexão, é suficiente e determinante para provocar o efeito desejado: a percepção da subida final, característica da pergunta.

• **comparação português/francês**

De forma inequívoca, a sobreposição dos dois esquemas melódicos permite-nos observar que os dois falantes utilizam duas estratégias muito diferentes, quando fazem uma pergunta: em francês, ataque bastante baixo e subida progressiva até um arranque final muito marcado; em português, ataque relativamente elevado e descida progressiva até ao momento de uma ligeira inversão do movimento, seguida de subida final.

Globalmente, temos praticamente duas curvas simétricas, relativamente à linha horizontal, passando pela frequência média das vogais situadas em meio de frase. Apesar das diferenças constatadas, existem pontos comuns: em início de frase, com um ataque ascendente nas duas línguas; em final, uma subida incidindo sobre a totalidade ou uma parte da última vogal.

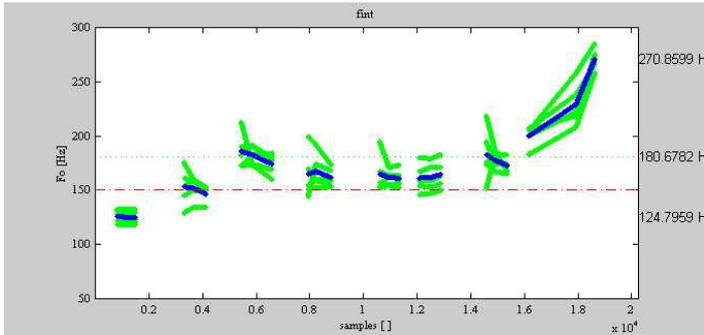


Figura 7 – Média das 5 repetições da frase: *Petitout passe à la télé* – francês – interrogativa

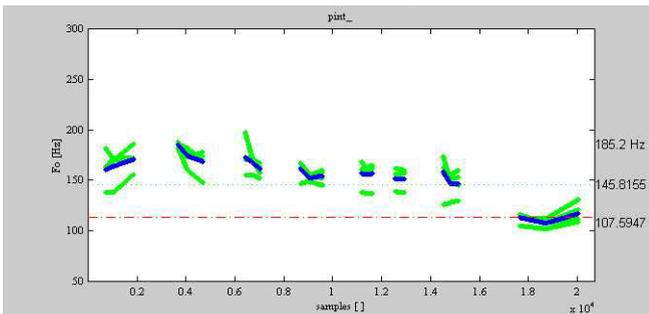


Figura 8 – Média das 5 repetições da frase: *Pateta toca no café* – português – interrogativa

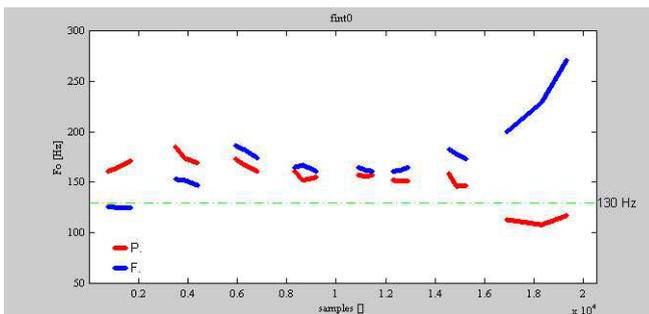


Figura 9 – Comparação das interrogativas PORT/FR

2.4. Variações intralocutores

Os gráficos que se seguem permitem comparar as variações de F_0 , para cada um dos falantes, não só entre as frases declarativa e a imperativa, mas também entre as frases declarativa e a interrogativa

Através da sobreposição da *declarativa-imperativa*, constata-se, em ambas as línguas, um comportamento bastante aproximado, apesar das diferenças individuais assinaladas. Na primeira parte da frase, F_0 sobe para a declarativa, enquanto que F_0 tem tendência para descer ou a inverter-se, na imperativa.

No segundo grupo rítmico, após um ataque sobre a primeira vogal da imperativa, a curva é descendente e praticamente sobreposta quando da realização das quatro últimas vogais.

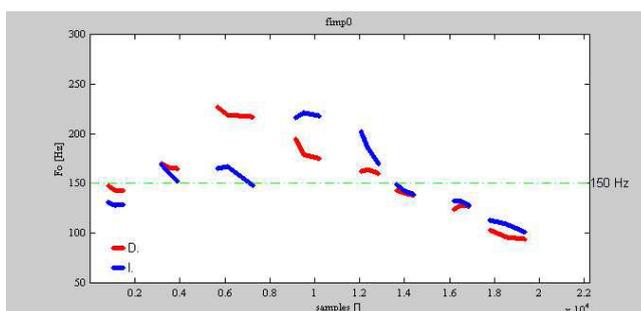


Figura 10 – Comparação entre a Declarativa e a Imperativa – falante francês
Frase: *Petitout passe à la télé*

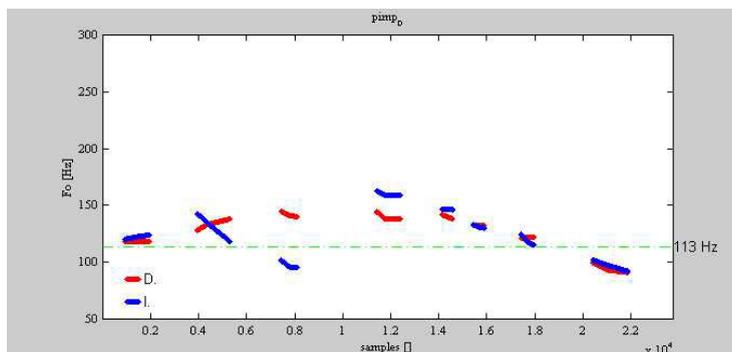


Figura 11 – Comparação entre a Declarativa e a Imperativa – falante português
Frase: *Pateta toca no café*

A sobreposição da *declarativa-interrogativa* confirma claramente as afirmações precedentes: a utilização de estratégias radicalmente diferentes nas duas línguas, relativamente à interrogativa:

- Em francês, o primeiro grupo rítmico sobe igualmente nos dois casos, mas com um valor da frequência fundamental ligeiramente mais elevado, no caso da declarativa (de 2 a 3 semitons sobre as três primeiras vogais). Este fenómeno pode talvez dever-se ao fato de que, neste caso, os dois elementos da frase são mais independentes do ponto de vista sintático, enquanto que, no caso da interrogativa, haveria uma maior coesão entre os grupos.
- Em Português, a curva melódica global dos dois tipos de frase é quase paralela, mas, desta vez, com um F_0 sistematicamente mais elevado em toda a frase interrogativa (um máximo de 5,6 semitons (118/164 Hz), no início; de 2,3 semitons (151/132 Hz) no resto da frase). A distinção entre os dois tipos de frase, no plano melódico, parece dever-se, essencialmente, a esta diferença de altura, sobretudo no ataque inicial e ainda a uma ligeira inversão ascendente sobre a última vogal, na frase interrogativa.

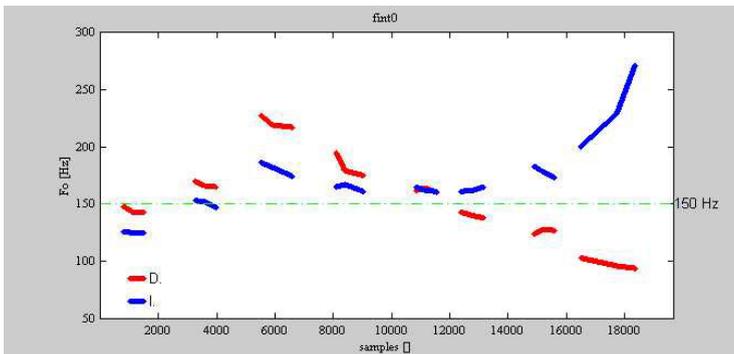


Figura 12 – Comparação entre a Declarativa e a Interrogativa – falante francês
Frase: *Petitout passe à la télé*

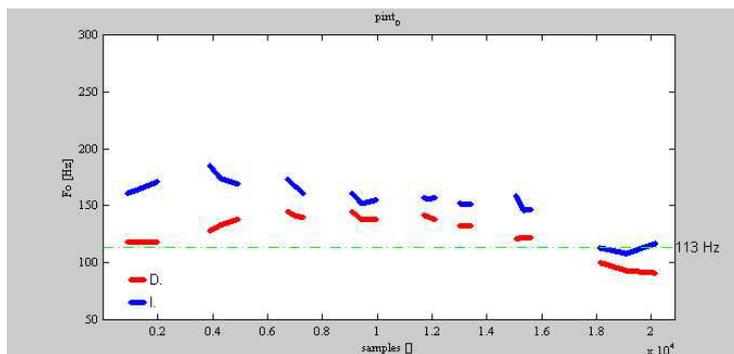


Figura 13 – Comparação entre a Declarativa e a Interrogativa – falante português
Frase: *Pateta toca no café*

Relativamente à realização entoacional dos dois *grupos rítmicos*, constata-se que, nas duas línguas, o primeiro grupo apresenta-se sempre de forma bastante distinta, com o ponto culminante sobre a sílaba acentuada: a segunda, tônica, em português; a terceira, rítmica, em francês. Observa-se também uma ligeira entorse à regra, no caso da declarativa em português, já que a terceira vogal átona de *Pateta* continua a ser produzida com F_0 elevado, fenômeno que pode ser considerado como um efeito acentual, mas sobretudo como um “eco” (com o sentido que lhe é atribuído por Delattre).

Uma ressalva deve ser colocada quanto às diferenças observadas, especialmente para as frases interrogativas. Com efeito, não é de todo impossível que estas diferenças, para além das diferentes estratégias utilizadas, possam refletir igualmente pequenas sutilezas na maneira de colocar a pergunta; por exemplo, uma pergunta retórica, do tipo “Pateta toca no café, não é verdade”?

3. Estudo da intensidade e da duração das vogais

3.1. A intensidade

É costume dizer-se que o parâmetro de intensidade desempenha um papel pouco importante em francês, quando comparado com

outras línguas, nomeadamente, com as línguas ditas acentuais. A observação dos gráficos que ilustram a variação deste parâmetro mostra que não há, efectivamente, nenhuma variação significativa da intensidade, o que permitiria detectar claramente os acentos ou os fenómenos demarcativos, independentemente do tipo de frase.

Os gráficos que se seguem, permitem observar, para o francês, a existência de um ponto comum: um decréscimo da intensidade (de 6 a 9 dB), sempre sobre a segunda vogal e duas vezes sobre a terceira vogal de *Petitout*, apesar da presença do acento rítmico na última sílaba [tu]. Esta constatação é um indício de que este acento será realizado por outros parâmetros, como, por exemplo, Fo. Esta diminuição dos valores da intensidade é, sem dúvida, imputável, pelo menos em parte, à natureza fechada das duas vogais [i] et [u], bem como ao seu ensurdecimento parcial, resultante da presença de uma consoante surda, enfatizado por um fenómeno de palatalização, no caso de [ti]. Em seguida, a intensidade retoma um nível bastante estável, variando apenas de 2 a 3 dB (101-104, 100-103, 103-105, segundo a frase). Em contrapartida, quase nunca se assiste a um decréscimo final da intensidade, facto que poderá parecer surpreendente, visto ser universalmente aceito que as vogais perdem progressivamente a energia no decorrer da realização de qualquer frase.

Em português, de forma invariável, a intensidade atinge um ligeiro ponto máximo sobre as vogais tónicas de *Pate ta* e de *to ca*, para, em seguida, decrescer de forma regular, sensivelmente sobre as quatro últimas sílabas da frase, de 4 a 12 dB (103-95, 104-92, 103-99, dependendo da frase). A vogal final, apesar de tónica, sofre sempre uma queda dos valores da intensidade (de 1 a 4 dB), relativamente à vogal precedente. Poderemos, por isso, dizer que neste locutor, as variações vão mais de encontro a uma certa lógica esperada, caracterizada por um reforço acentual e um enfraquecimento final (DELGADO-MARTINS, 2002).

No estádio em que se encontra esta pesquisa, embora estas diferenças nas duas línguas se nos apresentem estáveis, parece-nos, no entanto, difícil podermos afirmar se elas desempenham ou não um papel importante na produção e na percepção das frases realizadas.

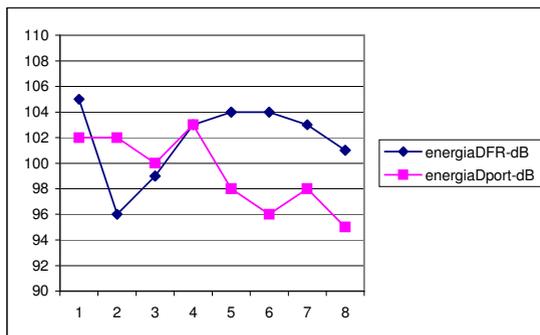


Figura 14 - Comparação da energia vocálica das Declarativas PT/FR

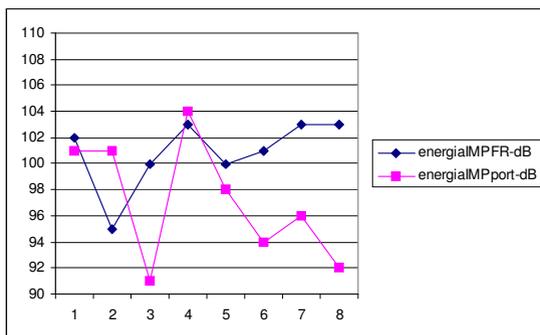


Figura 15 - Comparação da energia vocálica das Imperativas PT/FR

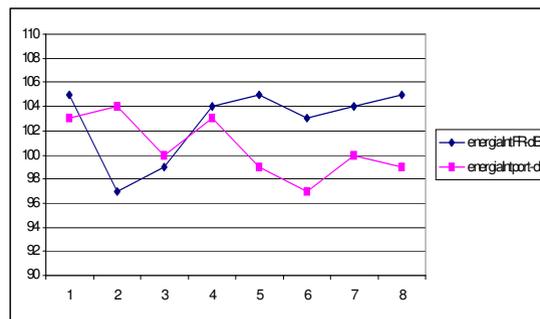


Figura 16 – Comparação da energia vocálica das Interrogativas PT/FR

3.2. A duração

O quadro seguinte, bem como os gráficos construídos com base nos mesmos valores, permitem-nos comparar as três frases nas duas línguas e observar o modo como é utilizado o parâmetro duração. Para além disso, oferece-nos ainda a possibilidade de podermos verificar se existem pontos comuns, quer entre os diferentes tipos de frase, quer entre as duas línguas.

No quadro, inscrevemos, com particular incidência, a duração das vogais que apresentam os valores de duração mais elevados, isto é, com uma duração geralmente superior a 40/60ms. No entanto, uma referência às vogais mais breves é também aí apresentada.

		1	2	3	4	5	6	7	8
tipo de frase	duração mínima	Pe	tit	tout	passé	à	la	té	lé
		Pa	te	ta	to	ca	no	ca	fé
Dec FR	39			64 (1,6)	63 (1,6)				81 (2)
Dec PT	31	66 (2,1)	65 (2,1)		57 (1,8)				103 (3,3)
Imp FR	40			136 (3,4)	70 (1,7)				113 (2,8)
Imp PT	28		108 (3,8)		69 (2,5)				82 (2,9)
Int FR	43			72 (1,7)	54 (1,3)				153 (3,6)
Int PT	26	71 (2,7)	62 (2,4)		56 (2,1)				150 (5,8)

Quadro 5 – Duração média (ms) das vogais mais longas em todas as frases (entre parênteses, a relação com a vogal mais breve da frase; a negrito as vogais nitidamente mais longas)

Constatamos, em primeiro lugar, que as durações vocálicas mais baixas atingem, aproximadamente, 40ms para o locutor francês, enquanto que para o português são de 30ms. Este facto pode parecer surpreendente, visto que o débito de fala do primeiro é ligeiramente superior ao deste último. Atribuímos esta particularidade ao facto de, em português, haver uma tendência para reduzir fortemente as vogais inacentuadas.

Em ambas as línguas, se considerarmos a posição na frase, observamos o seguinte:

- A *vogal final* é sempre longa ou muito longa (de 80 à 150 ms): entre 2 a 6 vezes superior às breves. Este facto explica-se certamente

pela sua posição no fim do primeiro grupo rítmico da frase. Como em português, para além da sua posição no grupo rítmico, também é portadora do acento tónico da última palavra (paroxítone), apresenta-se ainda mais longa do que em francês.

- As vogais *não finais*, esperadas como acentuadas ou potencialmente *acentuadas* (acento lexical em português, rítmico em francês), têm igualmente uma duração sempre ligeiramente superior: de 1,3 a 1,8 ms em francês e de 1,8 a 2,7 em português.

Ainda nas duas línguas, se compararmos os dois tipos de frase, constatamos que:

- A declarativa distingue-se essencialmente pela presença de uma vogal final longa (81 e 103 ms).
- A imperativa apresenta-se também com uma vogal final longa, com valores semelhantes aos obtidos para a declarativa: 82 e 113 ms. A vogal acentuada do 1º grupo é ainda mais longa, em ambas as línguas (108 e 136 ms). O sintagma correspondente ao vocativo é, deste modo, posto ainda mais em relevo do que propriamente a parte final da frase, que contém a ordem.
- Na interrogativa, o final é extremamente longo, para os dois informantes (150 e 153 ms, isto é, respectivamente 3,6 e 5,8 ms superior aos valores obtidos para as vogais breves).

Relativamente às vogais mais breves, o português realiza com maior duração do que o francês as vogais acentuadas. Não deixa de ser surpreendente, porquanto se diz que é o francês que privilegia o parâmetro duração.

Para além disso, o falante português utiliza as variações de alongamento vocálico (relação entre as longas e as breves) de forma superior às utilizadas pelo falante francês. É, sem dúvida, a sua forma de compensar a sua entoação mais monótona, já constatada, e à qual já tivemos ocasião de referirmos noutros momentos.

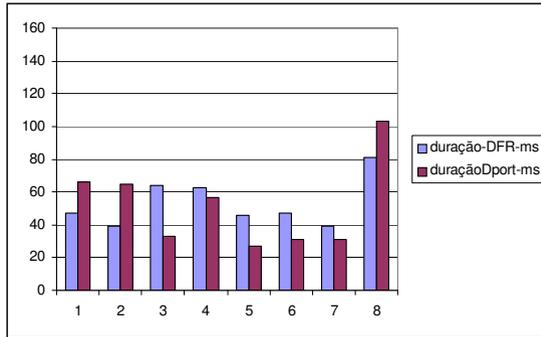


Figura 17 – Comparação das durações vocálicas das declarativas PORT/FR

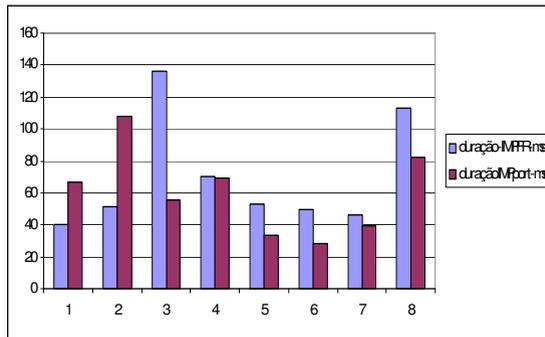


Figura 18 – Comparação das durações vocálicas das imperativas PT/FR

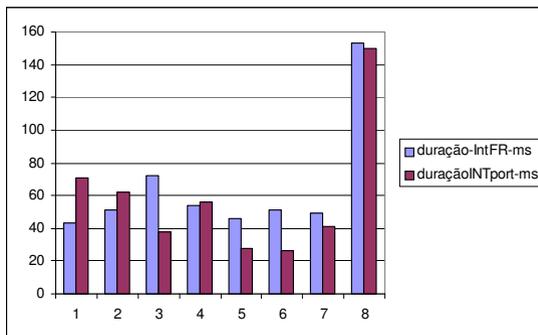


Figura 19 – Comparação das durações vocálicas das interrogativas PT/FR

SÍNTESE

Acabamos de verificar, e contrariamente a uma certa expectativa da nossa parte, que o comportamento prosódico das duas línguas, relativamente aos dois locutores estudados, é bastante próximo nos três tipos de frase estudados: declarativa, imperativa e interrogativa.

Para as duas primeiras, as diferenças são mínimas e dizem sobretudo respeito à diferença existente na posição do acento e ao diferente papel por ele desempenhado em cada uma das línguas. Os padrões melódicos e as variações de duração parecem bastante próximos em ambas as línguas.

A diferença mais notória que nós pudemos observar diz respeito às duas diferentes estratégias melódicas utilizadas para a frase interrogativa, embora permaneçam pontos comuns entre as duas línguas, sobretudo no início e fim de frase.

Fica-nos, no entanto, a dúvida sobre os matizes da mensagem que cada informante queria, efectivamente, transmitir quando da formulação da pergunta.

Constatámos que a intensidade não desempenha um papel significativo na distinção entre os três tipos de frases estudados. Parece mais sujeita à variação em português, aumentando ligeiramente para as vogais acentuadas, decrescendo em fim de frase, enquanto que para o francês permanece constante.

Contrariamente a uma ideia pré-concebida, que defende que o francês utiliza essencialmente o parâmetro duração para marcar os acentos, enquanto que as outras línguas utilizariam sobretudo a intensidade e a frequência fundamental (daí a denominação de “acento tônico”), nos casos analisados, parece ser o português que privilegia aquele parâmetro, sobretudo no caso das vogais acentuadas.

Recorde-se, no entanto, que a frase escolhida para o francês não era propícia ao efeito de alongamento vocálico que geralmente encontramos em certas sílabas fechadas.

Em contrapartida, as variações melódicas globais são claramente superiores no falante francês. Infelizmente, os dados de que dispomos são insuficientes para podermos decidir se trata de um fenómeno passível de ser generalizado às duas línguas estudadas ou

se deve ser imputado aos locutores. A continuação desta pesquisa é indispensável para clarificar este tipo de problemas.

Para concluir, o quadro seguinte propõe uma síntese das principais características prosódicas comuns às duas línguas, para ambos os informantes.

2 locutores	Declarativa	Imperativa	Interrogativa
ataque Fo	baixo	baixo	alto
Fo médio	–	–	+
melodia grupo 1 & grpo 2	subida /descida	subida-descida /descida	ligeira subida / forte subida ++(FR) / descida-subida (PT)
duração vocálica máxima	+ vogal final	++ vogal acentuada 1º gr. + vogal final	++ vogal final
intensidade		nada de significativo	

Quadro 6 – Principais características prosódicas comuns aos dois locutores

CONCLUSÃO

Como dissemos na introdução, o estudo proposto não é mais do que um esboço sobre os comportamentos prosódicos do francês e do português europeu, realizado de forma não exaustiva. As conclusões apresentadas devem, por isso, ser consideradas com prudência e não permitem, evidentemente, uma generalização, a menos que venham a ser validadas por outras pesquisas posteriores. Para além desse aprofundamento, gostaríamos também de dar continuidade a este trabalho, enriquecendo-o com outros esquemas prosódicos e completando esta comparação com outras línguas românicas, como o espanhol e o italiano. O alargamento de uma pesquisa neste domínio, não exclui, a longo prazo, a possibilidade de análise a outras línguas não românicas.

No plano da investigação fundamental, em fonética, os nossos resultados poderão ser comparados com os já obtidos no âmbito do Atlas das línguas românicas (projecto AMPER, já mencionado anteriormente), que tem como objecto de estudo as variedades regionais das línguas românicas.

No nível das aplicações práticas, e especialmente no domínio do ensino das línguas vivas, estes resultados deveriam permitir, desde já, sublinhar a importância que um aprendente deve dispensar à entoação e ao ritmo, quando se propõe falar uma língua diferente da sua língua materna (FLE, PLE, etc.).

NOTAS

¹ Projecto do Centre de Dialectologie de Grenoble - Université Stendhal, intitulado **AMPER** (Atlas **M**ultimédia **P**rosodique de l'**E**space **R**oman).

² Dada a implicação da equipe de Aveiro no projeto AMPER, e os programas de análise de que dispõe para este efeito, decidiu-se adotar uma metodologia de análise semelhante à utilizada no âmbito do referido projeto, especialmente no que diz respeito aos parâmetros estudados.

³ Note-se que nos gráficos que representam a sobreposição do movimento das curvas melódicas, o primeiro traçado serve de modelo de normalização, inclusive para a escala horizontal. Por isso, a duração dos segmentos correspondem à duração da frase francesa e não à portuguesa.

REFERÊNCIAS BIBLIOGRÁFICAS

CONTINI, M. Lai J.-P.; ROMANO, A.; ROULLET, S.; MOUTINHO, L. de C.; COIMBRA, R. L.; PEREIRA, B. U.; SECCA RUIVO, S. Un projet d'atlas multimédia prosodique de l'espace roman. *Proceedings of the International Conference Speech Prosody 2002*. Aix-en-Provence, 11-13 April 2002. p. 227-231.

DELGADO-MARTINS, M. R. (1982). *Sept études sur la perception accent et intonation du portugais*. Lisboa: INIC, 1986.

DELGADO-MARTINS, M. R. *Fonética do português. Trinta anos de investigação*. Lisboa: Edit. Caminho, 2002.

DELATTRE, P. *Studies in French and comparative phonetics*. La Haye: Mouton, 1966.

HIRST, D.; DI CRISTO, A. (Ed.). *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 1998.

MIRA MATEUS, M. H.; DUARTE, I. et al. *Gramática da língua portuguesa*. Coimbra: Almedina, 1982.

MORAIS BARBOSA, J. Entoação e prosódia. In: HOLTUS, G. et al. (Ed.). *Lexicon der Romanistischen Linguistik*, Band VI, n. 2, p. 243-248, 1994.

MOUTINHO, L. de C.; COIMBRA, R. L.; SECCA RUIVO, S.; PEREIRA, B. U. Atlas Prosódico Multimédia: Curvas de uma Trajectória. *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística*, Colibri, p. 387-391, 2001.

MARTIN, P. (1987). Pour une théorie de l'intonation. In: ROSSI et al. p. 234-271, 1981.

ROSSI. *L'intonation, le système français*. Orphis, 1999.

ANEXOS

1 – falante francês – declarativa – medidas médias para as 8 vogais

vogal	duração [ms]	energia [dB]	fo1	fo2	fo3 [Hz]
1	47	105	148	143	143
2	39	96	170	166	165
3	64	99	227	219	217
4	63	103	195	179	175
5	46	104	162	164	160
6	47	104	143	140	138
7	39	103	124	128	127
8	81	101	103	96	94

2 – falante francês – imperativa – medidas médias para as 8 vogais

vogal	duração [ms]	energia [dB]	fo1	fo2	fo3 [Hz]
1	40	102	131	128	129
2	51	95	169	160	152
3	136	100	165	167	148
4	70	103	216	221	218
5	53	100	202	186	169
6	50	101	149	142	139
7	46	103	132	132	128
8	113	103	113	109	101

3 – falante francês – interrogativa – medidas médias para as 8 vogais

vogal	duração [ms]	energia [dB]	fo1	fo2	fo3 [Hz]
1	43	105	126	125	125
2	51	97	153	152	147
3	72	99	186	182	174
4	54	104	165	167	161
5	46	105	165	162	161
6	51	103	161	162	165
7	49	104	183	178	173
8	153	105	200	229	271

4 – falante português – declarativa – medidas médias para as 8 vogais

vogal	duração [ms]	energia [dB]	fo1	fo2	fo3 [Hz]
1	66	102	118	118	118
2	65	102	128	133	138
3	33	100	145	141	140
4	57	103	145	138	138
5	27	98	142	140	138
6	31	96	132	132	132
7	31	98	121	122	122
8	103	95	100	93	91

5 – falante português – imperativa – medidas médias para as 8 vogais

vogal	duração [ms]	energia [dB]	fo1	fo2	fo3 [Hz]
1	67	101	120	122	124
2	108	101	143	135	118
3	56	91	102	96	95
4	69	104	163	159	159
5	33	98	147	147	146
6	28	94	133	131	130
7	39	96	125	118	115
8	82	92	102	97	92

6 – falante português – interrogativa – medidas médias para as 8 vogais

vogal	duração [ms]	energia [dB]	fo1	fo2	fo3 [Hz]
1	71	103	161	164	171
2	62	104	185	174	169
3	38	100	173	167	161
4	56	103	161	152	155
5	28	99	157	156	157
6	26	97	152	151	151
7	41	100	159	146	147
8	150	99	113	108	117

Intensidade e duração de formas reduzidas no português brasileiro

Larissa Santos Ciríaco
Bolsista do CNPq

Lorenzo Teixeira Vitral
Universidade Federal de Minas Gerais

César Reis
Universidade Federal de Minas Gerais

Abstract

In Brazilian Portuguese reduction occurs in some pronouns, for instance, *cê* (from *você*); *el* and *ez* (from the pronouns *ele* and *eles*); *num* (from *não*) and others that have been analyzed as becoming clitics (VITRAL & RAMOS, 2002). In order to prove this hypothesis, we collected a corpus with almost 500 utterances containing the reduced forms above. The reduced forms were acoustically analyzed considering the *intensity* and *duration* parameters and then compared with those parameters of real clitics. This paper presents the research results and the conclusions we have achieved in this work.

1. INTRODUÇÃO

Este trabalho visa investigar o comportamento acústico de algumas formas reduzidas do português brasileiro, a saber: *Cê* (redução de você), *Ez* (redução de eles) e *Num* (redução de não).

Em Vitral (1996), a primeira dessas partículas, isto é, *Cê*, foi avaliada como um clítico. O estatuto de clítico foi também estendido às partículas *Ez* e *Num* em Vitral (1999), Ramos (1997) e Corrêa (1998). A hipótese de Vitral (1996) está baseada no comportamento sintático de *Cê*, ou seja, a atribuição de clítico para as formas reduzidas é motivada por uma distribuição sintática dessas partículas. De acordo com esse autor, a partícula *Cê*, assim como os clíticos pronominais do português brasileiro, não pode ser encontrada nos seguintes ambientes: como tópico, como exemplificado em (1), modificados por um advérbio, como em (2), sem um hospedeiro no qual se possa ancorar, como em (3), como complemento de preposição, como em (4) e nem como elemento coordenado a uma forma tônica, como em (5). Os exemplos de (1) a (5) correspondentes à distribuição acima foram retirados de Vitral (1996) e são os que se seguem:

- 1)**Cê*, ele não viu.
- 2)*Só *cê* não consegue emprego.
- 3)*Eu amo *cê*. / Quem ganhou? **Cê*.
- 4)*Eu trouxe pra *cê*.
- 5)*Ele e *cê* podem votar contra.

A proposta de que *Cê* é um clítico é ainda corroborada pela análise de Vitral (no prelo), na qual são investigadas duas propriedades acústicas de *Cê*, a saber: a duração e a intensidade. Nesse trabalho, o autor mostra-nos que os valores da intensidade e da duração de *Cê* são menores do que os encontrados para o monossílabo *Zê*, e são

maiores do que os identificados para o clítico pronominal *Se*. Haveria, assim, uma gradação de valores para esses parâmetros, o que atestaria o caráter reduzido da forma *Cê*. Esses resultados levaram o autor a reformular sua hipótese inicial e a propor que *Cê* não é ainda um clítico pleno, mas se encontra em um *processo de cliticização*.

Neste trabalho, pretende-se desenvolver a metodologia proposta por Vitral, evitando certas deficiências encontradas, e estender sua perspectiva de análise para as demais formas reduzidas mencionadas, ou seja, *Ez* e *Num*. Aqui, trabalharemos, primeiramente, com a hipótese de que as reduções mencionadas podem ser atestadas foneticamente, e, posteriormente, que existe para elas um processo de cliticização e que esse processo pode ser constatado através da análise acústica.

O trabalho é estruturado da seguinte maneira: no item 2 forneceremos uma breve descrição do experimento; no item 3 falaremos dos parâmetros de análise acústica dos quais nos servimos e dos valores para esses parâmetros obtidos com a análise; a parte final trata das conclusões que o embasamento teórico e a análise nos permitem considerar.

2. OS PROCEDIMENTOS PARA A COLETA DE DADOS

Os dados deste estudo foram editados com o programa Sound Designer II e analisados através do programa de análise acústica Macquiner. Porém, antes da análise, estudamos uma maneira que permitisse a coleta adequada das formas reduzidas. Sabíamos que partir de dados de fala espontânea seria inviável, visto que a pesquisa tem por objetivo o comportamento fonético das partículas reduzidas, o que acarreta a observância do contexto fonético em que ocorrem as realizações dos segmentos. Por isso, desenvolvemos um experimento composto de treze sentenças cuidadosamente escolhidas, observando-se os contextos e as propriedades fonéticas de duração total, ritmo e acento das mesmas.

O experimento foi realizado com quatro informantes, sendo dois de cada sexo, todos estudantes do primeiro período do curso

de Letras da Universidade Federal de Minas Gerais, com idade entre 18 e 21 anos e nascidos e desde sempre habitantes da cidade de Belo Horizonte, Minas Gerais¹. As gravações foram feitas em cabine acusticamente tratada, no Laboratório de Fonética da Faculdade de Letras da UFMG e utilizando-se de gravador digital. Cada informante fez três leituras das treze sentenças sendo que seis das sentenças eram pronunciadas três vezes a cada leitura, ou seja, seis das sentenças foram realizadas nove vezes cada uma por cada informante em três velocidades de fala diferentes, a saber: normal, rápido e rapidíssimo. As sete sentenças restantes foram pronunciadas apenas uma vez a cada leitura, ou seja, três vezes cada uma por cada informante, sem a utilização do recurso da velocidade. Vale ressaltar que não houve comparação entre as partículas para as quais utilizou-se do recurso da velocidade e aquelas para as quais esse recurso não foi utilizado, a comparação deu-se entre formas plenas e respectivas formas reduzidas, como entre *Eles* e *Ez*, por exemplo. Desse modo, nosso *corpus* é composto por 75 realizações das sentenças por cada informante, totalizando 300 realizações.

O recurso de velocidade foi utilizado a fim de se obter a produção das formas reduzidas *Ez* e *Num*. As três velocidades de fala utilizadas eram normal, rápida e rapidíssima e não lenta, normal e rápida, como era de se supor, porque com isso obteríamos as reduções com segurança. Com esse recurso pudemos obter as formas reduzidas, mesmo tendo nas frases escritas as formas ortográficas *Eles* e *Não*. A forma reduzida *Ez* foi obtida principalmente na velocidade rapidíssima, ocorrendo também algumas vezes na velocidade rápida e outras poucas na velocidade normal. A redução *Num* ocorreu facilmente na velocidade rapidíssima e algumas vezes nas velocidades normal e rápida. Também fazia parte do conjunto das seis frases pronunciadas com o emprego da velocidade a forma *Ex*², (como em ex-marido ou ex-mulher). As sete sentenças restantes continham o clítico *Se* e a ocorrência de *CÊ* (grafado aqui completamente em caixa alta para distinguirmos-lo de *Cê*, redução de *você*) como referência a nome próprio. Para isso, nossos informantes foram instruídos de

que *CÊ* seria o apelido do nome *Cecília*. Assim, ao serem pronunciadas as sentenças que continham essa partícula, os informantes o faziam tendo em mente que estavam se referindo a um nome próprio e não a um pronome de tratamento, como quando pronunciavam o *Cê* enquanto forma reduzida de *você*. A redução de *você* foi escrita normalmente no experimento, em sua forma ortográfica *Cê*, por ser um segmento mais difundido no falar mineiro e por já não causar estranheza ao ser lido (em linguagem de internet, por exemplo, o uso de *Cê* escrito já é bem comum). Portanto, também fazia parte de nosso propósito a preocupação de que as realizações de nossas sentenças fossem o mais naturais possível, embora produzidas em laboratório. Para nos aproximarmos da fala espontânea, pedíamos que as frases fossem guardadas rapidamente na memória antes de serem pronunciadas, ou seja, nosso intuito era o de que o informante não ficasse preso à leitura. As sentenças do experimento são as que se seguem³:

1. Diga *Cê* disse que Zé vai viajar. (*Cê* =Cecília)
2. Diga *eles* falou que Zé vai viajar.
3. Diga *Jô não* disse nada não.
4. Diga *cê* disse que Zé vai viajar. (*cê* = pronome)
5. Diga *cê* falou que Zé vai viajar. (*cê* = pronome)
6. Diga o estilista vestiu *Cê* bem. (*Cê* =Cecília)
7. Diga *eles* disse que Zé vai viajar.
8. Diga o estilista vestiu-*se* bem.
9. Diga *Cê* falou que Zé vai viajar. (*Cê* =Cecília)
10. Diga *Jô não* falou nada não.
11. Diga o estilista vestiu *cê* bem. (*cê* = pronome)
12. Diga *ex* diz muita coisa.
13. Diga *eles* diz muita coisa.

Após a coleta dos dados e observados todos esses procedimentos acima citados, realizamos a edição e, posteriormente, a análise acústica dos segmentos em estudo.

3. ANÁLISE ACÚSTICA

A análise espectrográfica dos segmentos de nosso *corpus* nos permitiu medir os valores para os parâmetros de intensidade e duração. Utilizamos-nos destes parâmetros porque, em Moraes (1995) e em Fernandes (1976), foram estudados os correlatos acústicos do acento lexical, atestando a duração e a intensidade como marcas do acento. A duração foi considerada o parâmetro mais consistentemente presente como correlato do acento primário. Nossos estudos baseiam-se na hipótese, como dissemos, de que as formas reduzidas *Cê*, *Num* e *Ez*, devido aos indícios sintáticos, estariam em processo de cliticização. Sendo clíticos partículas átonas, buscamos assim investigar os valores de alguns desses parâmetros das marcas acentuais a fim de averiguar os indícios fonéticos desse processo de cliticização.

A análise é baseada nos valores obtidos das medições da *intensidade* e da *duração* das três formas reduzidas *Cê*, *Num* e *Ez*; das formas plenas *Eles* e *Não*; e da forma livre *Ex*. Também fazem parte de nossa análise o clítico *SE* e a realização de *CÊ* como nome próprio (apelido de Cecília). Após as medições, comparamos os valores de *CÊ* com os valores de *Cê* (redução de “você”) e de *Se* (clítico); os valores de *Não* e de *Num* e os valores de *Eles*, com os de *Ex* e de *Ez*.

Todas essas partículas acima ocupavam os seguintes ambientes nas sentenças: antes de sílaba tônica, isto é, antes de *disse* ou *diz*, como em 1, 3, 4, 7, 12 e 13; antes de sílaba átona, isto é, antes de *falou*, como em 2, 5, 9 e 10 e como elemento pós verbal⁴ para as formas *Cê* (pronome), *CÊ* (apelido de Cecília) e *Se* (pronome reflexivo), isto é, nas sentenças 6, 8 e 11.

Após o cálculo das médias das ocorrências das formas clíticas e plenas, nas três velocidades, no caso das formas para as quais esse recurso foi utilizado, os resultados obtidos para a *duração*, em milésimos de segundos (ms), e para a *intensidade*, em decibéis (dB) nos ambientes fonológicos descritos acima, são os mostrados na Tabela 1 abaixo. Essa tabela está assim representada: cada grupo corresponde a uma etapa de comparação. Assim, o grupo 1 refere-

se às realizações de *Eles*, *Ez* e *Ex*; o grupo 2 refere-se às realizações de *Não* e *Num*; o grupo 3 às realizações de *CÊ* e de *Cê* antes de verbo e o grupo 4, às realizações de *CÊ*, *Cê* e *Se* depois de verbo. Para cada grupo, temos os números das sentenças correspondentes, conforme numeração do experimento mostrado acima, um trecho da realização dessas sentenças e os respectivos valores para a duração e a intensidade. Ressalte-se que nem sempre as realizações das formas correspondem ao que está escrito nas sentenças, pois, como dito anteriormente, utilizamos alguns recursos para obter as reduções, ou seja, esperava-se que as formas *Eles* e *Não* fossem pronunciadas como *Ez* e *Num* nas velocidades rápida e rapidíssima, como explicado mais acima.

Médias				
Grupos	Nº Sentença	Realização	Duração (ms)	Intensidade (db)
1	02	...eles falou...	285.6	54.6
	02	...ez falou...	153.8	56.4
	7 e 13	...eles diz/disse...	287.1	56.7
	7 e 13	...ez diz/disse...	165.7	59.3
	12	...ex diz...	185.2	58.6
2	10	...não falou...	179.8	51.7
	10	...num falou...	101.7	55.9
	3	...não disse...	171.7	54.9
	3	...num disse...	128.9	56.5
3	9	...CE (Cecília) falou...	222.6	57.9
	5	...ce (pronome) falou...	152.5	55.9
	1	...CE (Cecília) disse...	271.7	56.1
	4	...ce (pronome) disse...	224.9	56.5
4	6	vestiu CE (Cecília)..	233.4	56.6
	11	vestiuCe (pronome)	204.9	54.9
	8	...vestiu-se....	132.9	54.9

Tabela 1: Médias da intensidade(db) e da duração(ms) para as partículas estudadas

3.1. Análise dos valores obtidos para a duração (ms)

Analisando os valores para a duração, os resultados vieram ao encontro da nossa hipótese de que podemos atestar, foneticamente, as formas reduzidas. É importante ressaltar que não tínhamos a pretensão de identificar essas formas reduzidas marcadamente como clíticos, pensamos apenas que existe, para estas partículas, um processo em andamento. Talvez elas sejam, por ora, apenas formas fracas, e pode ser que venham a incorporar características de clíticos verdadeiros. Enquanto isso, nosso estudo tenta identificar esse comportamento fonético das reduções do português brasileiro.

O que verificamos foi uma gradação da duração entre as partículas que usamos para a comparação. Observemos agora, mais detalhadamente, os resultados aferidos acerca da duração. A Tabela 2 abaixo, novamente organizada em grupos de sentenças para efeito de comparação entre elas, nos permite visualizar melhor a gradação dos valores entre as formas plenas e as reduzidas. Infelizmente, por nosso experimento ainda não se encontrar completo, restando algumas gravações a serem feitas, não poderemos estender a análise seguinte às realizações de *Eles*, *Ez* e *Ex*. Podemos apenas considerar aqui a gradação da duração para essas partículas de forma global, como mostrado na Tabela 1. No entanto, uma análise mais detalhada como a que se segue não será possível para essas palavras devido à falta de dados dessas partículas seguidas de sílabas átonas. Colocaremos os dados das partículas *Eles*, *Ez* e *Ex* seguidas de sílaba tônica apenas como dado adicional. Portanto, voltemos agora à análise mais detalhada das medidas para as partículas *Não*, *Num*, *CÊ*, *Cê*, e *Se*.

Grupo	Realização + Ambiente	Duração (ms)
1	Eles + sílaba tônica	287,1
	Ex + sílaba tônica	185,2
	Ez + sílaba tônica	165,7
2	Não + sílaba átona	179,8
	Num+ sílaba átona	101,7
3	Não + sílaba tônica	171,7
	Num + sílaba tônica	128,9
4	Ce (nome) + sílaba átona	222,6
	Ce (pronome) + sílaba átona	152,5
5	Ce (nome) + sílaba tônica	271,7
	Ce (pronome) + sílaba tônica	224,9

Tabela 2: Duração(ms) das partículas conforme ambiente fonético.

Como podemos observar, obtivemos uma gradação dos valores para todos os segmentos analisados e, em todos os cinco grupos, a forma plena *Não* e a partícula *CE*, significando nome próprio, mostraram uma duração maior do que suas respectivas formas reduzidas, estando sempre nos mesmos ambientes. Nos grupos 1 e 2, por se tratar de um dissílabo, é provável que o pronome *Eles* teria maior duração do que sua redução *Ez*, mas não se trata apenas de uma questão silábica; pois, como os dados apontam, a redução *Ez* possui, ainda assim, uma menor duração do que o também monossílabo *Ex*.

Outro importante fato observado foi a constância de uma maior diferença entre os valores medidos para a duração em ambiente “partícula + sílaba tônica”. Examinando-se os grupos 1, 2 e 3, 4 e 5, observa-se em todos uma gradação dos valores da duração atestando um processo de cliticização para as formas reduzidas. Mas, em 2 e 4, onde o ambiente é “partícula + sílaba átona”, a diferença de duração é maior do que em 3 e 5, cujo contexto é “partícula + sílaba tônica”. Embora as reduções também possuam menor duração ocorrendo no contexto das sentenças 2 e 4, esse valor é ligeiramente maior do que no contexto das sentenças 3 e 5. O quadro dessas diferenças em contexto tônico/átono é o seguinte: no grupo 2, onde

a sílaba seguinte à partícula é átona, o intervalo entre *Não* e *Num* foi aproximadamente 78 ms (diferença entre os valores 179.8 ms para *Não* e 101.7 ms para *Num*), enquanto para o grupo 3, onde a sílaba seguinte à partícula é tônica, a diferença foi de, aproximadamente, 43 ms. O mesmo ocorreu para os grupos 4 e 5. Em 4, a diferença é de 71 ms para a duração entre as formas plena e reduzida e de 47 ms em 5. Esse fenômeno ocorreu na totalidade de nossos dados.

Recorrendo à literatura, encontramos uma explicação condizente e possível para esse fenômeno, na estruturação rítmica do português. Sendo uma língua de ritmo acentual, é provável que o pé métrico seja uma unidade que exerça uma influência sobre a organização temporal dos enunciados. Observemos, por exemplo, as sentenças 3 e 10 (vide experimento no item 2 deste artigo) e suas duas realizações efetuadas, com a marcação dos pés métricos (entre barras) de cada uma mostrada abaixo:

Sentença 03: /Diga/Jô não /disse nada/ não.

Ou:/Diga/ Jô num /disse nada/ não.

Sentença 10: /Diga/Jô não fa/lou nada/ não.

Ou:/Diga/ Jô num fa/lou nada/não.

Essas duas sentenças estão aqui mostradas cada uma em suas duas realizações ocorridas. A sentença 3 foi realizada de duas formas: uma com a partícula estudada em sua forma plena, isto é, *Não*, e outra com a partícula em sua respectiva forma reduzida, isto é, *Num*. O mesmo pode ser dito para a sentença 10, mostrada logo abaixo da sentença 3. Para a análise que se segue, o segundo pé métrico ou a segunda é o que interessa.

Cada pé métrico das sentenças acima compreende uma sílaba tônica e as sílabas átonas que se seguem até a sílaba tônica seguinte, quando um novo pé métrico se inicia.

No estudo do ritmo, é conhecida a noção de isocronia, na qual os pés métricos tendem a ter a mesma duração, ou seja, essa é uma propriedade em que o tempo de realização dos pés métricos de uma frase tende a ter a mesma duração. Isso não quer dizer que um pé

com duas sílabas terá a mesma duração que um pé com uma ou três sílabas, mas que os valores para a duração dos mesmos tendem a se aproximar. Comparando-se o segundo pé métrico de cada uma das sentenças 3 e 10 entre si, em 10 o número de sílabas no segundo pé é maior, tanto para a ocorrência com forma reduzida quanto para a ocorrência com forma plena: a sentença 10 tem três sílabas métricas em seu segundo pé enquanto que a sentença 3 tem apenas duas. Tendo em vista o princípio de isocronia, os pés das sentenças tendem a ser proferidos no mesmo tempo. Conseqüentemente, o segundo pé métrico da sentença 10, que contém três sílabas, terá uma duração semelhante ao pé métrico da sentença 3, só que dividida pelas suas três sílabas. Isso provavelmente explica o fato de a duração das partículas reduzidas ser menor antes de sílaba átona do que antes de sílaba tônica. No entanto, é importante dizer que o princípio de isocronia funciona no interior de cada enunciado. Como as condições experimentais nos levam a crer que a organização temporal desses enunciados é a mesma, pensamos que a hipótese, neste estudo preliminar, do efeito do ritmo pode fornecer uma primeira explicação.

Em ambiente pós-verbal, as partículas *CÊ*, *Cê* e *Se* obtiveram as seguintes medições para a duração, como mostrado na tabela 3:

Ambiente	Partícula	Duração (ms)
Pós-Verbal	<i>CÊ</i> (nome)	233,4
	<i>Cê</i> (pronome)	204,9
	<i>Se</i> (reflexivo)	132,9

Tabela 3: Duração(ms) das partículas *CÊ*, *Cê* e *Se* em ambiente pós-verbal.

Aqui observamos, novamente, a gradação da duração entre a forma não clítica *CÊ*, a forma reduzida *Cê* e a forma clítica *Se*.

3.2. Análise dos valores obtidos para a Intensidade (dB)

Já a intensidade parece não distinguir as formas plenas das formas reduzidas da mesma maneira como obtivemos com o

parâmetro da duração. O esperado, seguindo a análise da duração, seriam valores menores aferidos para a intensidade das formas reduzidas do que para as formas plenas, aproximando as reduções ao pronome clítico *Se* e atestando o processo de cliticização proposto por Vitral. No entanto, o que ocorreu parece não corroborar esse processo de cliticização. Observou-se⁵ que a intensidade parece estar, na verdade, compensando a duração; numa comparação entre formas plenas e reduzidas. Vejamos novamente os valores para a intensidade, em contraste com os valores para a duração, das formas plenas e de suas respectivas formas reduzidas na tabela 4. Observe que retiramos as comparações entre *CÊ* (Cecília) e *Cê* (pronome), pois estas não são partículas concorrentes, ou seja, a primeira não corresponde à redução da segunda, tal qual *Eles* e *Ez*. Assim, a tabela 4 nos dá o contraste entre a duração e a intensidade numa comparação entre formas plenas e suas respectivas formas reduzidas.

Realização	Duração (ms)	Intensidade (dB)
...eles falou...	285.6	54.6
...ez falou...	153.8	56.4
...eles diz/disse...	287.1	56.7
...ez diz/disse...	165.7	59.3
...ex diz...	185.2	58.6
...não falou...	179.8	51.7
...num falou...	101.7	55.9
...não disse...	171.7	54.9
...num disse...	128.9	56.5

Tabela 4: Contraste entre duração e intensidade das formas concorrentes.

A partir dos dados da tabela 4, observamos o seguinte fenômeno: a intensidade parece compensar a duração na medida em que esses parâmetros se mostram em distribuição inversamente proporcional. Vejamos as partículas *Eles* e *Ez*; a duração da forma plena *Eles* é de 285.6 ms (antes de sílaba átona) e de 287.1 ms (antes de sílaba tônica); e sua intensidade é de 54.6 dB (antes de sílaba átona) e 56.7 dB (antes de sílaba tônica). Para sua forma reduzida *Ez*,

a duração cai para 153.8 ms (antes de sílaba átona) e para 165.7ms (antes de sílaba tônica) enquanto que a intensidade sobe para 56.4 dB (antes de sílaba átona) e para 59.3 dB (antes de sílaba tônica). Isto é, foneticamente, essa redução da duração é compensada na intensidade. O mesmo ocorre para *Não* e *Num*, basta observarmos a tabela acima.

Portanto, duração e intensidade apresentam comportamento diverso nos contextos forma plena/forma reduzida.

4. CONCLUSÕES

Iniciamos este trabalho com o objetivo de atestar, foneticamente, as formas reduzidas do português brasileiro *Ez*, *Num* e *Cê* e com a hipótese de que haveria, para estas reduções, um processo de cliticização, como proposto por Vitral.

Através da análise, vimos que esse processo de cliticização pode ser confirmado tanto pelo parâmetro da duração como da intensidade. Pudemos, assim com essa análise, atestar foneticamente as reduções de *Eles*, *Não* e *Você*.

Portanto, as reduções do português brasileiro aqui estudadas possuem as seguintes características fonéticas: há, para as partículas *Ez*, *Num* e *Cê* uma menor duração em comparação com as formas plenas *Eles* e *Não* e com a partícula correspondente a *Cê*, marcadamente não clítica, *CÊ* (significando nome próprio); e maior intensidade numa comparação apenas entre formas reduzidas e respectivas formas plenas (excetuando-se aí *CÊ* (Cecília) e *Cê* (você) pois não são partículas concorrentes tal qual *Eles* e *Ez* ou *Não* e *Num*), o que nos leva a concluir que talvez a queda constatada na duração das formas reduzidas estaria sendo compensada pela intensidade. Trata-se, no entanto, de estudo preliminar, que merecerá desenvolvimentos posteriores.

No entanto, reconhecemos que este trabalho conta com poucos dados numéricos para tirarmos maiores conclusões para o processo de cliticização proposto por Vitral. Por ora, pudemos atestar as formas reduzidas e elucidar alguns aspectos de seu comportamento fonético.

NOTAS

¹ Essas variáveis sócio-culturais não serão analisadas neste trabalho.

² Tivemos o cuidado de que a partícula *Ex* fosse assim realizada, i. e., não nos interessava, nem fazem parte deste estudo, realizações desta partícula como a forma ditongada *Eix*.

³ As sentenças que contêm o pronome *Eles* se encontram com o verbo no singular e o sujeito no plural de propósito. Este foi mais um recurso utilizado a fim de se obter a redução *Ez*.

⁴ Como dito na introdução, clíticos e, conseqüentemente, formas reduzidas não ocorrem como elemento pós-verbal numa proposição quando não possuem um hospedeiro no qual possam se ancorar. É importante frisar que isso não ocorre em nosso experimento, porque, embora na sentença 11, “Diga o estilista vestiu *Cê* bem”, por exemplo, a partícula *Cê* esteja em ambiente pós-verbal, seu hospedeiro é o advérbio *bem*. Segundo a sintaxe gerativa, a partícula *Cê* forma uma oração reduzida onde *Cê* é o sujeito e *bem* o predicado.

⁵ Observação feita pelo Prof. José Olímpio de Magalhães.

REFERÊNCIAS BIBLIOGRÁFICAS

BRANDÃO DE CARVALHO, J. Phonological conditions on Portuguese clitic placement: on syntactic evidence for stress and rhythmical patterns. *Linguistics* 27, p. 405-436, 1989.

CAGLIARI, L. C. *Algumas considerações sobre a duração silábica. Um estudo com dados do Projeto de Gramática do Português Falado*. Campinas: Unicamp. IEL, 1993.

CORRÊA, L. T. *A forma clítica de pronome pessoal no dialeto mineiro: uma variante sociolinguística*. 1998. Dissertação (Mestrado em Estudos Linguísticos) – Faculdade de Letras da UFMG.

DELGADO-MARTINS, M. R. *Ouvir falar. Introdução à fonética do português*. Caminho: Lisboa, 1988. Cap. 2, 3, 5 e 6.

FERNANDES, N. *Contribuição para uma análise instrumental da acentuação e intonação do português*. 1976. Dissertação (Mestrado) – USP, São Paulo.

JASSEM, W. *The phonology of modern English*. Warszawa: Panstwowe Wydawnictwo Naukowe, 1987.

JOHNSON, K. *Acoustic and auditory phonetics*. Malden, MA-USA: Blackwell Publishing, 2003.

MORAES, J. A. de. Acentuação lexical e acentuação frasal em português. Um estudo acústico-perceptivo. *Estudos Lingüísticos e Literários*, n. 17, Universidade Federal da Bahia, p.39-57, jul. 1995.

RAMOS, J. *O uso das formas Você, Ocê e Cê no dialeto mineiro*. In: DA HORA (Ed.). *Diversidade Lingüística no Brasil*, 1997. p.43-60.

VITRAL, L. T. A forma Cê e a noção de gramaticalização. *Revista de Estudos da Linguagem*, n. 5, p.115-124, 1996.

VITRAL, L. T. *Identifying clitics: phonetic evidence*. Artigo inédito, UFMG, 2001.

VITRAL, L. T.; RAMOS, J. Gramaticalização de “você”: um processo de perda de informação semântica? *Filologia e Lingüística Portuguesa*, Humanitas Publicações, São Paulo, v.3, p.55-64, 1999.

Tarefa de discriminação de fonemas com pseudopalavras*

Rui Rothe-Neves
Regina Carla Lapate
Juliana Sardinha Pinto
Universidade Federal de Minas Gerais

Abstract

This paper presents a task for phoneme discrimination as part of a more comprehensive set of tasks for psycholinguistic assessment of language currently under development (TAL-UFGM). The task consists of 40 pairs of monosyllabic nonwords, from which 20 pairs are made of two copies of the same syllable and 20 pairs differ in relation to one single feature (voice, place or mode of articulation) of their first consonant. This list of syllable pairs was presented in random order to 8 adult aphasic patients (experimental group) and 55 normal young adults (control group). Participants were asked to indicate after each pair whether it consisted of either equal or different syllables. The experimental group made 60.6% correct answers versus 97.5% by the control group. Differences were larger among voiced/unvoiced stimuli. The consistency of the task (Kuder-Richardson) for the experimental group was 0.823. Thus, the task is able to distinguish between normal and abnormal populations and, within each, between stimulus differences.

INTRODUÇÃO

Este artigo apresenta uma tarefa de discriminação de fonemas, parte integrante de um conjunto de tarefas de avaliação psicolinguística atualmente em desenvolvimento (TAL-UFMG). Na abordagem psicolinguística de avaliação (CAPLAN, 1992; ROTHE-NEVES & CAMARGOS, 2000), as tarefas são planejadas de modo a permitir a avaliação controlada dos sub-processos de produção e compreensão de linguagem. Portanto, não se trata de tarefas de avaliação de distúrbios ou patologias específicas de adultos ou crianças, mas tarefas que visam avaliar o funcionamento adequado dos diversos processos envolvidos na linguagem – a disfunção pode ser inferida a partir da comparação com o desempenho funcional, próximo a 100%. As tarefas comumente utilizadas no País para investigar o nível fonêmico da linguagem por meio da habilidade de discriminar fonemas, em geral, não cumprem esse requisito, pelos motivos discutidos a seguir.

O método mais utilizado para avaliar a discriminação de fonemas do português brasileiro (PB) em contexto clínico é o da repetição de palavras e pseudopalavras (ou logatomas), palavras inventadas que, embora sem sentido, obedecem às regras fonotáticas da língua (revisão em CHAVES; NEPOMUCENO; ROSSI; MOTA; PILLON, 1999; ver também DELGADO; BEVILACQUA, 1999; LONGONI; BORGES, 1997). Apresentam-se palavras e pseudopalavras ao sujeito, que deve repeti-las, e o escore da tarefa consiste na proporção de estímulos repetidos corretamente. A primeira observação crítica que se pode fazer em relação a esse método diz respeito ao comprometimento da repetição em pacientes com déficits articulatorios. O comportamento exigido como resposta não distingue erros provocados pelo prejuízo de natureza articulatória daqueles perceptivos. Se é desnecessário supor tais dificuldades em pacientes portadores de perda auditiva,

o mesmo não ocorre na afasia. Além do prejuízo à motricidade oral, que pode ser considerável, Varney (1984) estimou em 18% a proporção de pacientes afásicos com dificuldades na discriminação de fonemas. Da revisão da literatura oferecida em Chaves et. al. (op. cit., p.54) e Penrod (1989), depreende-se, ainda, duas limitações do método. Estímulos como palavras polissílabas ou frases inteiras oferecem pistas fonotáticas, prosódicas e semânticas, o que, por sua vez, podem servir à discriminação de fonemas. Além disso, ao se utilizarem palavras que de fato existem na língua, aparentemente supõe-se que a discriminação de fonemas não sofra influência de pistas semânticas. Superestima-se, assim, uma visão estritamente seqüencial do processamento lingüístico, a despeito da existência de propostas teóricas (ver revisão em ROTHE-NEVES; HAASE, 2000) e evidências experimentais em contrário (p.ex., MARSLEN-WILSON; TYLER, 1980). A partir dessas observações críticas, pode-se concluir que, embora o método da repetição possa servir *ad hoc*, não é adequado para uma avaliação da discriminação de fonemas que vise colher informações de grupos não-clínicos e clínicos de vários tipos.

Para superar tais dificuldades, a tarefa deve visar à avaliação da capacidade de distinguir traços característicos de fonemas, controlando variáveis externas, tais como a influência do *status* semântico, e dispensar o uso da linguagem na resposta. Desde Miller & Nicely (1955), sabe-se que a discriminação de consoantes oferece maior dificuldade quando se manipula apenas um traço distintivo. Na literatura psico-acústica, a mais simples avaliação comportamental da percepção é feita pelo procedimento de “escolha forçada de uma alternativa” ou procedimento “sim-não” (*yes-no procedure*). No contexto clínico, Caplan (1992) sugere uma tarefa com tais características. Apresentam-se ao sujeito 40 pares de pseudopalavras, sendo 20 de monossílabos idênticos e 20 distintos, em ordem aleatória. Os pares distintos diferem com relação a um único traço de um fonema consonantal (ponto de articulação, modo de articulação ou vozeamento). O fonema diferente pode ocorrer no início ou no final do estímulo e a consoante é seguida da vogal /a/. “As consoantes foram escolhidas como os segmentos a serem alterados porque têm sido usadas mais

freqüentemente do que vogais em pesquisa de discriminação de fonemas na afasia” (CAPLAN, 1992, p.408). Como resposta, a pessoa deve apenas informar se os dois itens são idênticos, o que pode ser feito por meio de qualquer sinal convencionado.

O escore é a proporção de respostas corretas. No presente trabalho, visamos construir e avaliar uma tarefa desse tipo com fonemas do PB.

MÉTODO

Materiais

A tarefa compreende 40 pares de sílabas sem sentido, sendo 20 pares compostos por duas sílabas idênticas e outros 20 pares, por duas sílabas que diferem apenas quanto a um único traço (ponto, modo ou vozeamento) nas consoantes iniciais. Para a escolha dos estímulos, sílabas fechadas foram compostas por permutação aleatória. Ao contrário do sugerido por Caplan, optamos por sílabas fechadas em que todas as vogais possam ocorrer, a fim de diversificar o contexto articulatorio, o que aumenta a validade ecológica da tarefa. Como as sílabas foram escolhidas aleatoriamente, a influência do contexto não é sistemática. Utilizaram-se todos os fonemas consonantais e vocálicos existentes e combinados adequadamente no PB. Excluíram-se, então, as sílabas existentes como palavras do léxico (p.ex.: [max]). Dentre as demais, selecionaram-se aleatoriamente 40 sílabas, das quais 20 foram duplicadas, formando os pares idênticos. Para as demais 20 sílabas, foram produzidos os pares respectivos, cuja consoante inicial difere num único traço. Os pares de sílabas assim produzidos foram lidos por uma profissional e registradas digitalmente (44,1 kHz, 16-bit).¹ O material foi editado a fim de apresentar 0,2 s entre as sílabas e 5 s entre cada par (cf. CAPLAN & UTMAN, 1994). Todos os pares foram então ordenados aleatoriamente.

Amostras

A fim de avaliar a aplicabilidade da tarefa, investigaram-se seus resultados em dois grupos de sujeitos. O grupo experimental (n=8)

consistiu de pacientes afásicos adultos selecionados, após consentimento livre e informado, dentre os indivíduos com lesão cerebral atendidos na Clínica de Fonoaudiologia do HC-UFMG. Os pacientes não receberam nenhum tipo de remuneração para colaborar com a pesquisa, mas os resultados foram disponibilizados aos profissionais que os assistem. O grupo controle (n=55) consistiu de jovens adultos, selecionados dentre estudantes de curso de graduação em Psicologia (UFMG), que receberam créditos por sua participação na pesquisa.

Procedimento

Os estímulos foram apresentados individualmente em campo aberto em sala silenciosa. Os sujeitos indicaram sua resposta verbalmente ou por escrito. Alternativamente, a pessoa poderia indicar sua resposta por meio de um cartão contendo duas figuras: a figura da direita apresentava a palavra IGUAL e três quadrados vermelhos, que representavam o conceito “igual”; a da esquerda, a palavra DIFERENTE e três figuras geométricas distintas (círculo, quadrado e triângulo), cada qual de uma cor, também representando o conceito a que se referiam. Essa alternativa foi utilizada quando as dificuldades motoras dos sujeitos do grupo experimental impediram-nos de responder das outras maneiras. Em todos os casos, a resposta era registrada pelo examinador imediatamente. Para cada sujeito registrou-se a proporção de acerto.

RESULTADOS E DISCUSSÃO

Como se poderia esperar, os resultados do grupo controle mostraram maior proporção de acerto do que os do grupo experimental (TAB. 1). A hipótese nula de que os resultados de cada grupo foram idênticos para erros e acertos foi testada por meio do teste χ^2 . O comportamento observado nos dados difere significativamente da hipótese nula no grupo experimental ($\chi^2=4,241$, $p<0,05$), mas principalmente no grupo controle ($\chi^2=10,935$, $p<0,001$). Nos pares de teste, também foi possível notar diferenças entre o traço manipulado (TAB. 2), sendo a mais significativa nos pares cuja consoante inicial se distingue por modo de articulação. Tanto no grupo experimental

quanto no controle, houve mais erros nos pares que se distinguem pelo vozeamento da consoante inicial. Também foi possível notar diferenças em cada par apresentado. Isto mostra que a tarefa, foco do estudo, permite de fato investigar diferenças entre grupos clínicos e não-clínicos.

TABELA 1 – Resposta dos grupos em relação ao tipo de estímulo.

ITENS	GRUPOS					
	Experimental (n = 8)			Controle (n = 55)		
	Acerto	Erro	Total	Acerto	Erro	Total
Distratores	106 (33,1%)	54 (16,9%)	160 (50%)	1085 (49,3%)	15 (0,7%)	1100 (50%)
Teste	88 (27,5%)	72 (22,5%)	160 (50%)	1061 (48,2%)	39 (1,8%)	1100 (50%)
TOTAL	194 (60,6%)	126 (39,4%)	320 (100%)	2146 (97,5%)	54 (2,5%)	2200 (100%)

TABELA 2 – Proporção de acerto dos grupos nos pares de teste

OPOSIÇÕES	GRUPOS		
	Experimental (n = 8)	Controle (n = 55)	
<i>Ponto</i>	v ʒ	0,875	0,982
	t k	1	1
	f ʃ	0,625	1
	b g	0,625	0,945
	f s	0,75	0,982
	m n	0,5	0,982
	b d	0,25	1
	<i>média</i>	<i>0,661</i>	<i>0,984</i>
<i>Modo</i>	d z	1	0,982
	n l	0,75	1
	d n	0,625	1
	t ʃ ʒ	0,5	0,982
	t s	0,625	0,982
	z n	0,5	1
	<i>média</i>	<i>0,667</i>	<i>0,991</i>
	<i>Voz</i>	ʃ ʒ	0,25
t d		0,25	0,6
s z		0,25	0,964
k g		0,5	1
p b		0,375	0,982
f v		0,375	1
t ʃ dʒ		0,375	0,927
<i>média</i>		<i>0,339</i>	<i>0,922</i>

Uma estatística relevante, nesse contexto em que se avalia uma tarefa, é o coeficiente de consistência, por meio do qual se pode conhecer a consistência entre itens ou a homogeneidade da tarefa. Uma alta consistência indica a confiabilidade da tarefa. Para o caso, o procedimento mais utilizado é o cálculo do coeficiente r_{tt} de Kuder-Richardson (cf. ANASTASI & URBINA, 2000, p.94), pois a tarefa aqui examinada exige respostas “sim-não” a cada item. A consistência da tarefa é satisfatória ($r_{tt}=.82$), quando examinadas as respostas do grupo experimental. O valor do coeficiente pode ser interpretado diretamente, como a proporção de variância devida à tarefa e não a erros sistemáticos. Isto é, 82% da variância nos resultados são devidos ao que se quer medir de fato. No entanto, o coeficiente r_{tt} deve ser tomado, aqui, com cautela, na medida em que é sensível ao número de itens, à quantidade de acertos e à variância e, com isso, ao tamanho da amostra. Enquanto o grupo experimental tem variância notável, a amostra é pequena. O grupo controle apresenta características opostas: bom número e variância desprezível, pois quase todos discriminaram corretamente os itens. A relação entre essas estatísticas e seu efeito pode ser vista ao compararmos o r_{tt} do grupo experimental, do grupo controle e, também, apenas para ilustração, de um grupo fictício formado por ambos os grupos juntos (TAB. 3). Esse último serve aqui para mostrar que, com maior variância e quantidade de acertos e sujeitos, o r_{tt} seria, provavelmente, melhor – uma questão a ser futuramente investigada.

TABELA 3 – Consistência (r_{tt} de KUDER-RICHARDSON) e estatísticas relevantes

GRUPO	ESTATÍSTICA		
	Variância	Σpq	r_{tt}
Experimental	39,928	7,875	0,823
Controle	1,796	0,8	0,569
Ambos	97,249	2,466	0,99963

Um resultado interessante, que suscita novas investigações, é a diferença entre a proporção de acertos nos itens de teste e os distratores. Tecnicamente, o foco recai sobre os itens de teste, feitos

para possibilitar a distinção entre pessoas com e sem problemas de discriminação. Este objetivo, como se vê, é de fato alcançado pelos itens de teste. No entanto, o grupo experimental também se distinguiu do grupo controle nos itens distratores. Isto pode ser explicado pelo fato de que os distratores ofereceram dificuldade irrisória para o grupo controle, mas o grupo clínico encontrou mesmo aí uma dificuldade maior. Numa futura investigação, será necessário investigar a estrutura de respostas, incluindo, além dos acertos, também os falsos positivos e os falsos negativos, a fim de avaliar a sensibilidade e a especificidade de cada item.

CONCLUSÕES

A partir dos resultados apresentados, conclui-se que a tarefa de discriminação de fonemas aqui investigada pode ser útil na avaliação dessa capacidade, na medida em que seu procedimento mostrou-se compreensível, adequado às condições clínicas e produziu resultados confiáveis. Do ponto de vista empírico, novos estudos devem comparar a discriminação avaliada nessa e noutra tarefa, de modo a estabelecer estatisticamente sua validade, além da sensibilidade e especificidade dos itens. Também será útil investigar a aplicabilidade da tarefa no contexto do desenvolvimento normal e clínico da linguagem.

NOTAS

* Este trabalho foi apresentado numa forma preliminar em comunicação ao 7º Congresso Brasileiro e 1º Congresso Internacional de Fonética e Fonologia, Belo Horizonte, 22 a 30/10/2002. É parte integrante do projeto “Testes de avaliação de distúrbios de linguagem: a abordagem psicolinguística”, que contou com financiamento da PRPq/UFMG (convênio 5173*24) e bolsas PIBIC/CNPq (109788/2002-8 e 106786/2003-2).

¹ Nesta etapa colaboraram Maurílio Nunes Vieira (Laboratório de Voz, ICEx/UFMG) e Rachel Ferreira Loiola (Unicentro Izabela Hendrix e Programa de Pós-Graduação em Estudos Linguísticos, POSLIN/UFMG), a quem muito agradecemos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANASTASI, A.; URBINA, S. *Testagem psicológica*. 7.ed. Porto Alegre: Artes Médicas, 2000.
- CAPLAN, D. *Language: structure, processing, and disorders*. Cambridge, MA: MIT, 1992.
- CAPLAN, D.; UTMAN, J. A. Selective acoustic phonetic impairment and lexical access in an aphasic patient. *Journal of the Acoustical Society of America*, v.95, n.1, p. 512-517, jan. 1994.
- CHAVES, A. D.; NEPOMUCENO, L. A.; ROSSI, A. G.; MOTA, H. B.; PILLON, L. Reconhecimento de fala: uma descrição de resultados obtidos em função do número de sílabas dos estímulos. *Pró-Fono*, v.11, n.1, p.53-58, 1999.
- DELGADO, E. M. C.; BEVILACQUA, M. C. Lista de palavras como procedimento de avaliação da percepção dos sons da fala para crianças deficientes auditivas. *Pró-Fono*, v.11, n.1, p.59-64, 1999.
- LONGONI, E.; BORGES, A. C. C. Teste de reconhecimento de fala em indivíduos portadores de perda auditiva neurossensorial. *Pró-Fono*, v.9, n.1, p.3-9, 1997.
- MARSLEN-WILSON, W.; TYLER, L. K. The temporal structure of spoken language understanding. *Cognition*, v.8, p.1-71, 1980.
- MILLER, G. A.; NICELY, P. E. An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America*, v.27, p.338-352, 1955.
- ROTHER-NEVES, R.; CAMARGOS, L. Avaliação de distúrbios de linguagem: a abordagem psicolinguística In: 8ª Conferência Internacional de Avaliação Psicológica, 2000, Belo Horizonte. *Anais / 5 Encontro Mineiro de Avaliação Psicológica/8 Conferência Internacional de Avaliação Psicológica*. São Paulo: Vetor Editora Psicopedagógica, 2002. p.128-130.
- ROTHER-NEVES, R.; HAASE, V. G. O “sistema nervoso conceitual” e suas implicações para a avaliação neuropsicológica. II: paradigmas emergentes. *Cadernos de Psicologia da Sociedade Brasileira de Psicologia*, Ribeirão Preto, v.6, 2000.
- PENROD, J. P. Logaudiometria. In: KATZ, J. (Org.) *Tratado de audiologia clínica*. São Paulo: Manole, 1989. Cap.10, p. 146-162.
- VARNEY, N. L. Phonemic imperception in aphasia. *Brain and Language*, v.21, p.85-94.

Reorganização prosódica na disartria

Erica Reviglio Iliovitz
UNICAMP

Abstract

According to Felizatti (1998), dysarthria is a disorder in motor production that affects patterns of movement, coordination and strength of articulatory organs. It can be caused by brain damage and it is usually characterized by slow speech. The analysis of some prosodic aspects of almost unintelligible dysarthric patients' speech indicates that the strange reaction their speech causes in their listeners is mainly due to the prosodic strategies they use (consciously or not) to improve their oral language intelligibility. In other words, despite the fact that dysarthric patients's brain lesions cause articulatory difficulties, they are able to compensate them through prosody.

0. INTRODUÇÃO

Neste texto, serão apresentados alguns aspectos prosódicos da fala geralmente lenta e pouco inteligível de sujeitos com disartria pós-traumática. Esses aspectos foram constatados através de um estudo longitudinal (i.e., realizado no decorrer de um determinado período de tempo).

Um dos objetivos dos trabalhos longitudinais em geral é o de analisar e comparar as mudanças ocorridas nos aspectos estudados ao longo do tempo. No trabalho longitudinal realizado com dois sujeitos disártricos (RV e LG), no período de um ano (agosto/2001 a agosto/2002), têm sido verificadas certas alterações prosódicas que podem indicar em que medida a fala deles foi irremediavelmente alterada em decorrência das lesões cerebrais e em que medida os sujeitos, ao se depararem com o estranhamento próprio e alheio em relação à própria fala, se utilizaram de elementos prosódicos já presentes no sistema da língua materna para tentarem torná-la mais compreensível para os ouvintes.

Uma das hipóteses é que o recurso a domínios prosódicos superiores (cf. SCARPA, 2000; no prelo) se deu em função de dificuldades de articulação¹ decorrentes dos danos cerebrais apresentadas pelos sujeitos. Numa análise preliminar dos dados - juntamente com relatos de entrevistas - pudemos constatar que, embora ambos os sujeitos recorram a elementos prosódicos para compensar as dificuldades articulatórias no nível segmental, eles fizeram opções diferentes nesse sentido. Enquanto a fala de RV (que apresenta um quadro disártrico mais grave) demonstra grandes variações melódicas, a de LG apresenta poucas variações. Além disso, conforme disse em entrevista, ele contraía propositalmente suas pregas vocais, e isso tornava sua voz mais grave. A justificativa

dada por ele para tal atitude é que, ao fazer isso, ele “se lembrava” de que tinha que “mexer mais a boca” para falar, i.e., que ele tinha que articular melhor os sons. RV, por outro lado, produzia grandes variações melódicas sem se dar conta disso, i.e., inconscientemente.

Nesse sentido, um dos objetivos do trabalho é mostrar que o estranhamento que a fala desses sujeitos com disartria pós-traumática provoca nos ouvintes se deve não só à lentificação da velocidade de fala e às mudanças de qualidade de voz² decorrentes das lesões cerebrais, mas, também, e principalmente, a essas estratégias prosódicas das quais eles se utilizam para viabilizar a inteligibilidade da própria linguagem oral. Em outras palavras, pretendo mostrar que os sujeitos cérebro-lesados em geral – e os disártricos em particular – se utilizam (ainda que intuitivamente, sem plena consciência disso) de determinados recursos prosódicos já disponíveis na língua para tentar tornar a própria fala mais compreensível – o que acaba promovendo uma reorganização prosódica na fala desses sujeitos.

Para apresentar essa reorganização prosódica, têm sido realizadas análises fonético-acústicas dos dados desses sujeitos no momento em que eles chegaram ao CCA (Centro de Convivência de Afásicos do Instituto de Estudos da Linguagem da UNICAMP) em 2001, para compará-las com os dados colhidos em 2002. Além disso, pretendo sistematizar, de acordo com os conceitos de prosódia presentes na literatura lingüística, as características da fala de RV (que, numa análise preliminar, apresentou uma subida de F_0 nas sílabas tônicas) e na qualidade de voz mais baixa de LG (que apresentou *pitch contour*, curva de altura, mais baixa comparada ao sujeito-controle).

Na primeira parte, serão apresentadas algumas hipóteses referentes às estratégias lingüísticas utilizadas pelos sujeitos disártricos. Na segunda parte, serão apresentados alguns conceitos referentes a estudos sobre a relação entre prosódia e disartria. Na terceira, a metodologia; na quarta, serão apresentados os sujeitos e suas falas; na quinta, as estratégias lingüísticas, e na sexta parte, as conclusões preliminares.

1. HIPÓTESES

As hipóteses aqui apresentadas se referem ao uso de elementos supra-segmentais, i.e., prosódicos, para compensar dificuldades articulatórias no nível segmental. Dos três elementos que compõem a prosódia, serão analisados dois: a duração³ e a altura.⁴

Em relação à duração, a hipótese é que a lentificação da velocidade de fala é compensada pela redução silábica, i.e., os sujeitos procuram enfatizar a articulação das sílabas tônicas em detrimento das átonas na tentativa de falar mais rapidamente; e, no que se refere à altura, as variações visam contribuir para a inteligibilidade da fala, tanto através da redução, quanto através do aumento de F_0 .

2. PROSÓDIA E DISARTRIA

Segundo Felizatti (1998), a disartria corresponde a uma desordem na produção motora que afeta os padrões de movimento, precisão, coordenação e força dos órgãos fono-articulatórios. Além disso, envolve lesões motoras de origem geralmente traumática no sistema nervoso central, em níveis cerebelares e subcorticais, configurando comprometimentos fonético-fonológicos causados pelo enfraquecimento dos músculos fonatórios.

De modo geral, estudos relativos à prosódia e disartria concluíram que há lentificação generalizada da velocidade de fala, desarranjo rítmico causado pela neutralização da duração silábica, maximização do uso de pausas e não-redução das vogais átonas finais (cf. SCARPA, 2000). Desta forma, a alternância rítmica e o âmbito limitado da curva de altura (distância de F_0 entre níveis altos e baixos dos contornos entonacionais) caracterizam alguns problemas de comunicação expressos pelos sujeitos disártricos que, contudo, exibem preservação de níveis da hierarquia prosódica, pois inserem pausas entre fronteiras prosódicas maiores (cf. SCARPA, 2000; no prelo).

3. METODOLOGIA

A fala dos sujeitos disártricos LG e RV foi gravada em um gravador digital portátil – *mini-disc* da marca SONY (modelo MZ-R 700/R700 PC/ R700 DPC), em sessões fonoaudiológicas, no CCA (Centro de Convivência de Afásicos) do IEL-UNICAMP (Instituto de Estudos da Linguagem da Universidade Estadual de Campinas). As gravações foram feitas por mim durante três períodos de seis meses.

A primeira gravação ocorreu em 08/2001, quando foram gravados trechos de fala espontânea (relato de situação cômica) e leitura de um texto-controle; em 02/2002, gravamos uma entrevista na qual os sujeitos relatam as principais queixas em relação à própria fala e leitura do mesmo texto-controle; finalmente, em 08/2002, foi gravada uma conversa na qual os sujeitos falam sobre esportes e lêem, novamente, o mesmo texto.

Para análise acústica, as gravações foram segmentadas em orações e, em seguida, os espectrogramas foram analisados no programa PRAAT, desenvolvido por Paul Boersma, do Instituto de Ciências Fonéticas da Universidade de Amsterdam.

Um sujeito controle, i.e., sem lesão cerebral (TT) também realizou a leitura do texto com o propósito de fornecer um parâmetro comparativo no nível segmental. TT é paulista como LG e RV; nasceu em Jundiaí no dia 20/10/1978, é destro e musicista/músico, com escolaridade superior completa.

O texto lido foi “História Estranha”, de Luiz Fernando Veríssimo. A escolha desse texto foi baseada na curta extensão (21 orações) e no fato de todas as orações serem declarativas, o que permite a manutenção de um padrão entonacional. De acordo com Moraes 2000: 183,

em português, como na maioria das línguas conhecidas, o padrão da declarativa é caracterizado por uma queda de F_0 no fim da pronúncia (mais precisamente, na tônica final), enquanto o *pitch* inicial está no nível médio.

4. OS SUJEITOS DISÁRTRICOS

4.1. LG

LG nasceu em 23/05/1970 em São Paulo/SP. Em outubro de 1998, aos 28 anos, sofreu um acidente automobilístico. Em exames imediatamente pós-acidente, foram observadas lesões corticais e subcorticais caracterizadas por áreas hiperdensas (edemas) em região frontal esquerda, cápsula nuclear direita e região parietal alta direita.

Ficou internado durante 40 dias e em coma induzido por 10 dias. Teve traumatismo craniano à direita e, como seqüelas, uma lesão cerebelar relacionada aos aspectos motores da face e da fala, contusão cerebral e lesão axional difusa grave. Três meses após o acidente, apresentou euforia e fala inarticulada. Além disso, teve anacusia (perda da audição) à direita. O eletroencefalograma e o mapeamento cerebral permanecem normais. Houve um aumento das partes moles na região parietal direita e há presença de áreas hiperdensas na região frontal esquerda, cápsula nuclear direita e região parietal alta direita, com hipodensidade adjacente.

LG é destro e, atualmente, aos 32 anos, trabalha como engenheiro civil.

4.1.1. A Fala de LG

As características gerais da fala de LG são:

- achatamento de F_0 ; ex: (ver figuras 1 e 2 no final do texto);
- qualidade de voz: *breathy voice*;
- impressão auditiva geral: monotônico.⁵

4.2. RV

Nascido em 16/03/1973, em Campinas/SP, RV sofreu um acidente automobilístico em 12/07/1998 aos 25 anos de idade. Ficou em coma induzido durante 15 dias e teve traumatismo craniano à direita e lesão axional difusa grave, lesões corticais no lobo temporal

esquerdo com redução cortical, na pirâmide direita, ponte e pedúnculo cerebral direito, além do corpo caloso. Fez traqueostomia por 1 ano e, em decorrência de estenose, teve uma retirada cirúrgica de anéis traqueais com subsequente inserção de prótese de silicone. Voltou a falar três meses após o acidente. Em fevereiro de 1999, iniciou terapia fonoaudiológica e ecoterapia. Apresenta ausência de sensibilidade na porção anterior média e posterior do lado direito da boca e face.

RV é canhoto e, atualmente, aos 29 anos, estuda Direito.

4.2.1. A Fala de RV

A fala de RV apresenta as seguintes características:

- prolongamento e maximização do uso de pausas; tom alto nas sílabas tônicas; mudança de tessitura para faixa alta; ex: (ver figuras 1 e 3 no final do texto)
- qualidade de voz: *creaky voice*⁶ + *falsetto*⁷;
- impressão auditiva geral: hipermelódico.⁸

5. ESTRATÉGIAS LINGÜÍSTICAS

Tanto LG quanto RV desenvolveram estratégias lingüísticas para lidar com a própria fala após o acidente. Por exemplo, ao ser questionado a respeito do tom vocal baixo, LG disse que força intencionalmente as pregas vocais para se obrigar a articular melhor os sons. Desta forma, ele cria uma dificuldade articulatória (tensão voluntária e excessiva das pregas vocais) para tentar compensar outra, que envolve a produção de elementos segmentais (sílabas, palavras, frases).

A brevidade silábica característica da fala de LG também pode ser justificada como uma estratégia que visa compensar a lentificação da velocidade de fala, i.e., como uma tentativa de normalizar a duração.

Já RV, ao ser informado que falava com grandes variações melódicas e entonacionais, particularmente nas sílabas tônicas,

explicou que isso foi um recurso inconsciente do qual ele se utiliza para manter a atenção dos interlocutores – que, de uma maneira geral, tendiam a se aborrecer com a fala lenta dele. Assim, para compensar dificuldades de articulação no nível segmental, RV se vale de recursos suprasegmentais presentes na língua, intensificando-os.

Além disso, segundo ele o tom da voz dele ficou muito baixo/grave após o acidente, e as variações melódicas e entonacionais pareciam contribuir para atenuar essa impressão.

6. CONCLUSÕES PRELIMINARES

Devido às dificuldades de articulação decorrentes das lesões cerebrais, tanto LG quanto RV se utilizaram, conscientemente ou não, dos elementos prosódicos já presentes e disponíveis na língua para tentar melhorar a inteligibilidade da própria fala. Tal fato pode ser constatado a partir da análise fonético-acústica da fala dos sujeitos, que mostra alterações prosódicas em relação a um padrão.

NOTAS

¹ Tais dificuldades envolvem enrijecimento e fadiga muscular, o que perturba a articulação apropriada de fonemas, especialmente de consoantes (tepe e grupos consonantais) e a velocidade de fala. A maximização do uso de pausas e a lentificação da velocidade de fala ocorrem devido às dificuldades de controle respiratório dos sujeitos, que têm que inspirar mais vezes para administrar o fluxo de ar destinado à fala e à respiração propriamente dita.

² “Aspecto auditivo geral de uma determinada voz, com a contribuição de fatores laringais e supralaringais”. (LAVÉ, 1980). A qualidade de voz deriva de duas fontes diferentes: das características anatômicas do aparato vocal do falante e do uso que o falante faz do aparato vocal. “A noção de qualidade de voz é básica para o estabelecimento de qualquer tipo de estrutura conceitual que apresente a gama completa de contrastes lingüísticos e prosódicos. (...) Ela corresponde a uma combinação de parâmetros acústicos e articulatórios que variam de forma independente, sendo que os atributos mais importantes são altura, volume, duração e timbre”. (CRYSTAL, 1969: 123)

³ A *duração* é “a quantidade de tempo consumida por um evento discursivo; geralmente é expressa em milésimos de segundo (msg). É um traço fonético.” (cf. LAVER, 1994: 431). Pode ser vista como a duração de tempo que um falante decide continuar a produzir uma unidade lingüística, a duração dos correlatos acústicos da unidade no espectrograma ou como a duração de tempo na qual o ouvinte ouve a unidade. (cf. CRUTTENDEN, 1986/1992).

⁴ A *altura* é a variável do pico de voz em uma sílaba ou em um número de sílabas sucessivas. Fisiologicamente, depende da taxa de vibração das pregas vocais dentro da laringe. A taxa de vibração das pregas vocais (correlato articulatório) é refletida na medição acústica da *Frequência Fundamental* (F_0). Em outras palavras, F_0 é o correlato acústico (físico) da altura. (cf. CRUTTENDEN, 1986/1992).

⁵ Tom de voz que apresenta âmbitos limitados da curva de altura, i.e., baixas variações de F_0 .

⁶ A voz do tipo *creak* é caracterizada pelo fornecimento de *input* com pulsos de energia para o trato vocal. Esses pulsos, porém, ocorrem numa frequência muito baixa e são espaçados irregularmente no tempo (cf. LAVER, 1994: 194-195), o que gera vibração irregular das pregas vocais.

⁷ No *falsetto*, as pregas vocais são esticadas longitudinalmente de frente para trás, tornando-se relativamente finas na seção cruzada, com as margens glotais das pregas ficando particularmente finas. A glote fica ligeiramente aberta. Assim, o *falsetto* é acompanhado por um leve sussuro, já que o fluxo de ar pulmonar consegue escapar continuamente pela passagem glotal. Além disso, a pressão subglotal é caracteristicamente mais baixa do que na fonação vozeada comum, em parte por causa do contínuo vozeamento transglotal (LAVER, 1994: 197).

⁸ Que apresenta variação excessivamente alta da altura (cf. COLSHER, COOPER & GRAFF-RADFORD (1987:379)).

REFERÊNCIAS BIBLIOGRÁFICAS

COLSHER, P. L.; COOPER, W.E.; GRAFF-RADFORD, N. Intonational variability in the speech of right-hemisphere damaged patients. *Brain and Language* 32, p. 379-383, 1987.

CRUTTENDEN, A. *Intonation*. Cambridge: Cambridge University Press, 1986/1992.

CRYSTAL, D. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press, 1969.

FELLIZATTI, P. *Aspectos fonético-fonológicos da disartria pós-traumática: um estudo de caso*. 1998. Dissertação (Mestrado) – IEL/UNICAMP.

LAVÉ, J. *Principles of phonetics*. Cambridge: Cambridge University Press, 1994.

LAVÉ, J. *The phonetic description of voice quality*. Cambridge: Cambridge University Press, 1980.

MORAES, J. A. Intonation in Brazilian Portuguese. In: HIRST, D.; DI CRISTO, A. (Ed.). *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 2000. p.179-194.

SCARPA, E. M. Abordagens prosódicas à linguagem de sujeitos cérebro-lesados. *GEL/2000*.

SCARPA, E. M. Dificuldades prosódicas em sujeitos cérebro-lesados. *Revista ALFA* (no prelo).

VERISSIMO, L. F. História Estranha. In: *Comédias para se ler na escola*. São Paulo: Ed. Objetiva, 2000. p.47.

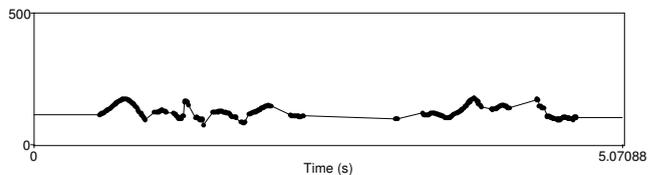
FIGURAS 1, 2 E 3**A) Figuras 1 e 2**

Figura 1 - TT: "O homem tenta dizer alguma coisa, mas não encontra o que dizer."

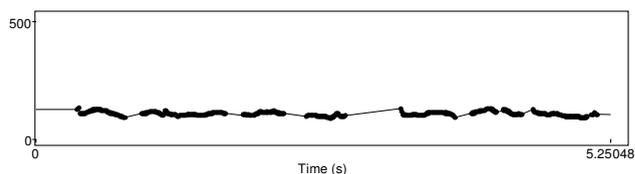


Figura 2 - LG: "O homem tenta dizer alguma coisa, mas não encontra o que dizer."

Leitura da frase "O homem tenta dizer alguma coisa, mas não encontra o que dizer", feita pelo sujeito controle TT e pelo paciente disártrico LG. O F_0 de LG é achatado em relação ao do controle. Além disso, a pausa entre a primeira e segunda orações é reduzida.

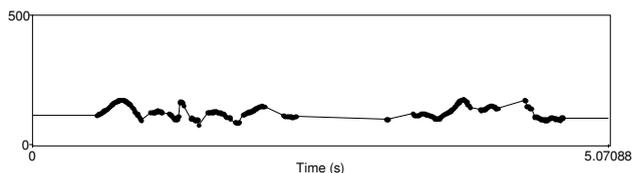
B) Figuras 1 e 3

Figura 1 - TT: "O homem tenta dizer alguma coisa, mas não encontra o que dizer."

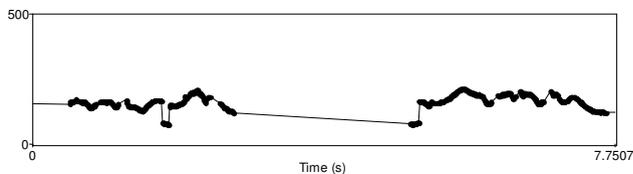


Figura 3 - RV: "O homem tenta dizer alguma coisa, mas não encontra o que dizer."

Leitura da frase "O homem tenta dizer alguma coisa, mas não encontra o que dizer", feita pelo sujeito controle TT e pelo paciente disártrico RV. O F_0 de RV é hipermelódico em relação ao do controle. Além disso, a pausa é prolongada.

A pausa na produção da fala com comprometimento neurológico

Jussara Melo Vieira
Plínio Almeida Barbosa
Maria Inês Pegoraro-Krook
Universidade de São Paulo

Abstract

Dysarthria is associated with a deficit in phonemic planning or paralysis of specific speech articulators, besides a frequent velopharyngeal impairment and expected prosodic alterations. The main goal of this work is to describe the speech rhythm of a user of a palatal lift protheses by analyzing the way she signals some linguistic boundaries when reading a short story. Pause durations were measured and pauses were classified into six syntactic categories. The correlation coefficient between categories and durations is 62%, signaling that not all variance can be solely explained by syntax, but also by considering prosody as the interplay between linguistic knowledge and constraints of a biomechanical system. Despite her neuromotor impairment, the subject is able to use pauses to signal a prosodic hierarchy.

INTRODUÇÃO

A fala é um fenômeno cultural, exclusivamente humano, somente possível por sua vida em sociedade (SAPIR, *apud* DUBOIS et al., 1998).

Transmitir um pensamento é muito mais que concatenar fonemas. Sempre se recorre a parâmetros que estão além da estrutura segmental, nos planos fonético/fonológico, sintático/semântico da língua falada, como a prosódia, que lhes confere um colorido peculiar. Esse colorido pode ser, entre outros, a expressão de sentimentos, emoções e/ou idéias do falante.

A prosódia é um supra-segmento que possui dois aspectos: 1) produção, identificado pelos três parâmetros clássicos: a duração – representada pela diferença de tempo entre dois eventos, a frequência fundamental e a intensidade; 2) percepção, identificada pelas noções de duração percebida, altura e volume (BARBOSA, 1999).

Como a produção da prosódia está inserida na fala, como esta para que seja adequada, é necessária a integração do sistema nervoso central, da audição, do sistema respiratório, da laringe e das cavidades supraglóticas (PEGORARO-KROOK, 1995).

Do ponto de vista fisiológico-acústico, a fala tem origem na vibração das pregas vocais, provocada pelo ar expirado dos pulmões, produzindo um som (voz) que será modelado pelos articuladores (lábios, bochechas, mandíbula, dentes, palato duro, palato mole, paredes faríngeas) e ressonadores (regiões supraglóticas da laringe e faringe, cavidades oral e nasal), que compõem o trato vocal. O comprometimento de quaisquer uma dessas estruturas pode resultar em prejuízo para a fala.

Os indivíduos com lesões neurológicas, congênitas ou adquiridas, podem apresentar disartria que é uma desordem neuromuscular com fraqueza, incoordenação, levando a uma perturbação da articulação dos fonemas devido ou a uma lesão central (lesão cortical), que causa

um déficit na programação desses fonemas ou a lesões periféricas, que causam a paralisia de certos órgãos motores de execução da atividade de fala (DUBOIS et al., 1998). O tipo de disartria relaciona-se ao local e à extensão da lesão cerebral (BOONE & PLANTE, 1994). Os músculos da língua, dos lábios e do palato mole podem estar afetados e apresentarem incoordenação, fraqueza, hipertonia ou hipotonia e a velocidade de fala pode estar diminuída conforme o grau de paresias e paralisias (ZIEGLER et al., 1988).

Pelo comprometimento neuromuscular, espera-se uma dificuldade em controlar a respiração, especialmente o fluxo expiratório no momento de fala, pois uma fadiga aparece conforme prolonga-se a fala, corroborando para uma incoordenação pneumo-fono-articulatória.

Também é freqüente a presença da incompetência velofaríngea (IVF) levando a uma ressonância hipernasal, pois o acoplamento entre as cavidades oral e nasal resulta em uma perda indesejada de ar (emissão de ar nasal) e de som durante a produção da fala, o que acarreta uma ressonância nasal excessiva, popularmente conhecida como fala fanhosa (PEGORARO-KROOK, 1995).

Problemas dessa natureza, não sendo tratados cirurgicamente, apresentam indicação para uso da prótese de palato elevadora, que consiste em um aparelho intra-oral removível com uma extensão (porção elevadora) feita de resina, a qual tem a finalidade de elevar o palato mole em direção à parede posterior da faringe (BÓ & PEGORARO-KROOK, 2001). Portanto, o uso da prótese de palato elevadora é o tratamento mais indicado para a IVF, quando associada à presença de paralisias envolvendo os músculos do esfíncter velofaríngeo (TACHIMURA et al., 2000), a fim de estimular a função velofaríngea em associação com a terapia de fala (TACHIMURA et al., 2001).

Diante desses comprometimentos neurológicos que afetam a fala, tratados ou não proteticamente, a prosódia está alterada.

Assim, o objetivo do presente trabalho é descrever, em termos de produção da estruturação rítmica, a fala de uma jovem neurologicamente prejudicada que faz uso da prótese de palato elevadora, para determinar como ela assinala fronteiras lingüísticas ao longo dos enunciados de uma narrativa.

METODOLOGIA

A participante desse estudo é a jovem I.M.M, regularmente matriculada no Hospital de Reabilitação de Anomalias Craniofaciais (HRAC) – USP/Bauru, do sexo feminino, com 26 anos e 3 meses de idade, na data de gravação (12/12/00), cujo comprometimento neurológico é em virtude de um traumatismo crânio-encefálico ocasionado por uma queda de uma mobilete. Por apresentar IVF, faz uso da prótese de palato elevadora desde 1995.

No Laboratório de Fonética Acústica do HRAC, sala tratada acusticamente, solicitou-se que a jovem, usando a prótese de palato elevadora, realizasse a leitura do texto exclusivamente composto por fonemas orais “A história do urso preto” (SUGUIMOTO & PEGORARO-KROOK, 1996) (Anexo 1). Essa leitura foi gravada em uma fita digital Sony, utilizando-se o *Digital Audio Tape (DAT) DTC-690*, também da marca Sony. Posteriormente, editou-se essa gravação via computador PC-IBM, por meio do programa WaveStudio (Creative Sound). A edição foi armazenada em disquete para *zip drive* e, então, analisada pelo *Multi-Speech*, modelo 3700, da Kay Elemetrics, da seguinte forma: trecho a trecho da leitura, com auxílio do espectrograma, fizeram-se marcações de duração de pausa (em milissegundos). As fronteiras das pausas foram assim determinadas: a primeira marcação correspondeu ao início da última vogal do segmento emitido antes da pausa, e a marcação final correspondeu ao início da primeira vogal do segmento seguinte (Fig. 1). A razão de uma tal marcação de pausa fundamenta-se nos trabalhos de Grosjean e colegas (1979, 1983), pois melhor revela a hierarquia das forças prosódicas subjacentes, o que é confirmado em Português Brasileiro (PB) no trabalho de Barbosa (1999), que ressalta, além disso, a importância do *onset* da vogal como delimitador por excelência de unidades prosódicas mínimas.



Figura 1 – Demonstração das marcações sobre os trechos da leitura do texto oral “A história do urso preto”

Após as marcações, estabeleceram-se categorias (Tab. 1), com base léxico-gramatical, para se caracterizar as fronteiras lingüísticas onde tais pausas aconteceram, na tentativa de se hierarquizar suas ocorrências. Quanto menor o número da categoria, maior a força da fronteira sintática por esse número representada e, conseqüentemente, maior o efeito da projeção dessa força sobre a cadeia da fala.

Tabela 1 – Descrição das categorias estabelecidas para esse estudo

Categorias	Tipo de fronteira lingüística
1	entre frases ((.)).
2	entre orações, após introdutores e no início de conjunções.
3	entre sintagmas nominal e verbal.
4	no sintagma verbal, entre sintagma nominal e sintagma preposicional; ou entre o verbo e sintagma nominal ou preposicional.
5	no início de sintagma preposicional dentro de sintagma nominal.
6	no início de sintagma preposicional dentro de sintagma preposicional dentro de sintagma adverbial.

De posse das durações e suas categorias, por meio do aplicativo Excel (Microsoft, 2001), estabeleceram-se as relações entre elas e aplicou-se a análise estatística de correlação, sabendo-se que se a taxa de correlação for próxima ou maior que 70%, mais da metade da variância dos dados pode ser explicada apenas pela assinalação de força de fronteira sintática via duração.

RESULTADOS E DISCUSSÃO

Na Tabela 2, é possível verificar as durações encontradas, em ordem crescente, e as categorias estabelecidas.

Tabela 2 – Duração (ms), fronteiras lingüísticas e suas categorias, a partir da leitura do texto “A história do urso preto”

Duração (ms)	Fronteiras Lingüísticas	Categorias	Localização no Texto*
62	zoológico da	6	F1
271	dias ele	4	B7
419	vive atrás	3	E2
498	dia de	5	D4
516	aparece logo	3	E1
543	barulho das	5	D3
557	história do	5	D2
560	grades do	5	D1
563	dia, resolveu	2	C5
566	preso. Decidiu	1	A7
580	cedo para	4	B6
584	olhar os	4	B5
597	floresta. Hoje	1	A6
600	ver os	4	B4
601	pássaros e	2	C4
614	descobriu que	4	B3
627	viu que	4	B2
628	jaula. Logo	1	A5
635	triste, pois	2	C3
636	Decidiu que	2	C2
637	preto que	2	C1
649	esperar o	4	B1
704	araras. Outro	1	A4
750	bichos. Gosta	1	A3
768	difícil. Ficou	1	A2
903	cidade. Todos	1	A1

* Em termos da Duração (ms): A>B>C>D>E>F; A1>A2>A3 e assim por diante.

As categorias que mais ocorreram foram as de número 1 e 4, com 7 ocorrências; seguidas das categorias 2, 5, 3 e 6 com 5, 4, 2, e 1 ocorrências, respectivamente.

Relacionando-se as categorias encontradas com as durações, obteve-se o Gráfico 1.

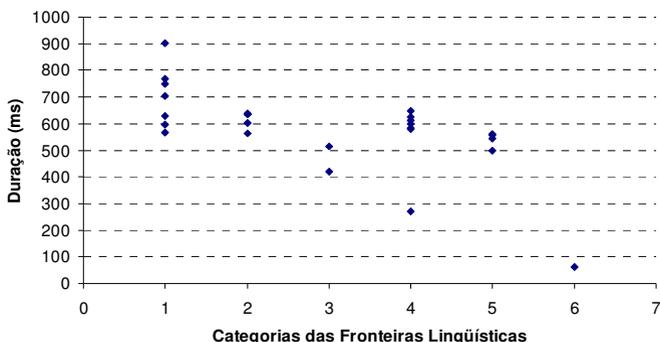


Gráfico 1 – Relação das categorias estabelecidas das fronteiras lingüísticas com sua duração, a partir da leitura do texto oral “A história do urso preto”.

Analisando-se o gráfico, percebeu-se que a concentração na categoria 1 ocorre com maior freqüência nas maiores durações, delineando-se uma tendência de quanto mais fraca a categoria, menor a duração.

Na análise estatística de correlação, obteve-se o valor absoluto de 0,62, o que corresponde a 62%. Esta correlação é significativa, porém menor que 70%. Isso assinala que nem toda a variância da duração pode ser explicada apenas pelas categorias sintáticas tais como escolhidas. Isso significa que elas precisam ser reavaliadas, notadamente com teorias sintáticas mais apropriadas, e que durações assinalam uma hierarquia prosódica que não é isomórfica com a hierarquia sintática, visto que a prosódia é o resultado da negociação entre um sistema lingüístico e um sistema biomecânico de produção de fala (BARBOSA, 1999).

A localização das pausas e sua duração possibilitaram o estabelecimento de sua relação com a saída de ar e a conseqüente fadiga respiratória, utilizando ar residual até o esgotamento do fôlego (FERREIRA, 1996), esperada em casos com este tipo de comprometimento neurológico. Ou seja, maior duração das pausas quanto mais esteja no final do texto lido.

Pela existência de pausas e pelo estabelecimento de suas durações em relação às fronteiras lingüísticas, percebeu-se que, apesar de seus comprometimentos neuromotores, a jovem é capaz de fazer uso das pausas para estabelecer uma hierarquia prosódica. Percebe-se, então, que o estabelecimento das pausas tem mais relação com a estrutura da sentença (GROSJEAN, GROSJEAN e LANE, 1979) do que com a condição pneumo-fono-articulatória do falante, apesar de se esperar desobediência à pontuação de um texto pela incoordenação pneumo-fono-articulatória, como afirma Ferreira (1996, p. 11).

CONCLUSÃO

Para o estabelecimento do quadro lingüístico de um falante é preciso ter muito claro até que ponto a fisiologia responde por seu desempenho lingüístico, fazendo, da melhor forma possível, sua relação com os aspectos puramente lingüísticos, avaliando a qualidade e não apenas a quantidade de sua competência lingüística. Espera-se que aspectos lingüísticos como a prosódia apresentem-se deficitários, em nível de desempenho, mas podem estar íntegros do ponto de vista da competência. Perspectivas deste tipo merecem atenção, para que se conheça, cada vez mais, os prejuízos de fala de indivíduos com comprometimento neurológico; sua capacidade lingüística remanescente; as estratégias compensatórias de que utiliza; o quanto a parte neuromotora responde e corresponde à sua competência lingüística; como estão os outros aspectos prosódicos, como o ritmo, a entonação e a acentuação, visto que as pausas ocorreram. A partir disso, poderão ser propostos procedimentos terapêuticos mantenedores ou ampliadores das atuais habilidades lingüísticas, da participante deste estudo bem como avaliar o uso da prótese de palato elevadora com finalidade fisiológico-acústico-lingüística. Assim, serão propostos a este falante uma melhor qualidade de vida e uma maior compreensão de sua comunidade lingüística quanto à sua competência/desempenho lingüísticos. Sem esquecer que o aprofundamento na estrutura da língua se faz necessário, a fim de se proporem categorizações cada vez mais adequadas e, sem esquecer, ainda, que a percepção prosódica do ouvinte, quanto ao desempenho do falante, também deve ser investigada.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARBOSA, P. A. Revelar a estrutura rítmica de uma língua construindo máquinas falantes: pela integração de ciência e tecnologia de fala. In: SCARPA, E. M. (Org.). *Estudos de Prosódia*, Campinas/SP: Editora da Unicamp, 1999. 327p. Cap. 1.
- BELL-BERTI, F.; GELFER, C.E.; BOYLE, M. Utterance-final lengthening: the effect of speaking rate. *ICPhS 95*, Stockholm, 1, sessão 8.3, p.162-5, 1995.
- BÓ, F. R.; PEGORARO-KROOK, M. I. Prótese de Palato. Disponível em: <<<http://www.fob.usp.br/depart/baf/protesedepalato/oque.htm>>>. Acesso em: 11/12/2001.
- BOONE, D.R.; McFARLANE, S.C. *A voz e a terapia vocal*. 5. ed. Trad. Sandra Costa. Porto Alegre: Artes Médicas, 1994. 300p.
- CRYSTAL, T.H.; HOUSE, A.S. Segmental durations in connected speech signals: preliminary results. *Journal of the Acoustical Society of America*, n. 72, v. 3, p.705-16, set. 1982.
- DUBOIS, J.; GIACOMO, M.; GUESPIN, L.; MARCELLESI, C.; MARCELLESI, J.B.; MEVEL, J-P. *Dicionário de Lingüística*, 10. ed. São Paulo: Editora Cultrix, 1998.
- FANT, G.; KRUCKENBERG, A. Quantal theory of speech timing. *Fonetik 96*, Swedish Phonetics Conference, Näslingen, p. 41-44, maio 1996.
- FERREIRA, L.P. Respiração: tipo, capacidade e coordenação pneumo-fono-articulatória. In: FERREIRA et al. *Temas de Fonoaudiologia*. São Paulo: Loyola, 1996. Cap.1, p.11.
- FOWLER, C. Production and perception of coarticulation among stressed and unstressed vowels, *Journal of Speech and Hearing Research*, n. 47, p. 127-139, 1981, *apud* BARBOSA, P.A. Revelar a estrutura rítmica de uma língua construindo máquinas falantes: pela integração de ciência e tecnologia de fala. In: SCARPA, E.M. (Org.). *Estudos de Prosódia*. Campinas/SP: Editora da Unicamp, 1999. 327p. Cap. 1.
- GROSJEAN, F. How long is the sentence? Prediction and prosody in the on-line processing of language. *Linguistics*, n. 21, p.501-529, 1983.
- GROSJEAN, F.; GROSJEAN, L.; LANE, H. The patterns of silence: performance structures in sentence production. *Cognitive Psychology*, n. 11, p.58-81, 1979.
- MONNIN, P.; GROSJEAN, F. Mémoires originaux. *L'Année Psychologique*, n. 93, p.9-30, 1993.

NOOTEBOOM, S.G. How far do we look ahead while speaking? *ICPhS 95*, Stockholm, 4, sessão 86.2, p.578-81, 1995.

PEGORARO-KROOK, M.I. *Avaliação da fala de pacientes que apresentam inadequação velofaríngea e que utilizam prótese de palato*. 1995. 130p. Tese (Doutorado) – Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo.

SAPIR, E. *apud* DUBOIS, J.; GIACOMO, M.; GUESPIN, L.; MARCELLESI, C.; MARCELLESI, J.B.; MEVEL, J-P. *Dicionário de Linguística*, 10. ed. São Paulo: Editora Cultrix, 1998.

SCOTT, D. R. Duration as a cue to the perception of a phrase boundary a), *J. Acoust. Soc. Am.*, n. 71, v. 4, p. 996-1007, abr. 1982.

STREETER, L. A. Acoustic determinants of phrase boundary perception. *J. Acoust. Soc. Am.*, n. 64, v. 6, dez. 1978.

SUGUIMOTO, M.L.C.P.; PEGORARO-KROOK, M.I. Avaliação nasométrica em adultos normais falantes do português brasileiro. *Pró-Fono: Revista de Atualização Científica*, n. 8, v. 1, p. 7-12, 1996.

TACHIMURA, T.; NOHARA, K.; HARA, H.; WADA, T. Effect of placement of a speech appliance on levator veli palatini muscle activity during speech. *Cleft Palate-Craniofacial Journal*, v. 37, n. 5, p. 478-482, set. 2000.

TACHIMURA, T.; NOHARA, K.; FUJITA, Y.; HARA, H.; WADA, T. Change in levator veli palatini muscle activity of normal speakers in association with elevation of the velum using an experimental palatal lift prosthesis. *Cleft Palate-Craniofacial Journal*, v. 38, n. 5, p. 449-454, set. 2001.

ZIEGLER, W.; HOOLE, P.; HARTMANN, E. D. Von Cramon Accelerated speech in dysarthria after acquired brain injury: Acoustic correlates. *British Journal of Disorders of Communication*, n. 23, p. 215-228, 1988.

ANEXO 1**A história do urso preto**

D2 C1 A6 E2
 Esta é a história do urso preto que veio da floresta. Hoje ele vive atrás
 D1 F1 A1 B7 E1
 das grades do zoológico da cidade. Todos os dias ele aparece logo
 B6 B4 A3 B5 C4
 cedo para ver os outros bichos. Gosta de olhar os pássaros e de ouvir
 D3 A4 C5 A5 B2
 o barulho das araras. Outro dia, resolveu sair da jaula. Logo viu que
 A2 C3 B3 A7 C2
 era difícil. Ficou triste, pois descobriu que estava preso. Decidiu que
 B1 D4
 deveria esperar o dia de sua verdadeira liberdade.

Banco de dados orais: uma nova perspectiva aos estudos do português brasileiro

Regina Célia Fernandes Cruz
Universidade Federal do Pará

Jailma do Socorro Uchôa Bulhões
Bolsista – IC – UFPA

Léa da Silva Fernandes
Bolsista – IC – UFPA

Abstract

This work represents an attempt to facilitate the automatic access to the spontaneous Portuguese data for studies on experimental phonetic area and speech synthesis, maintaining qualitative and quantitative information of the data. This corpus constitutes more than 30 hours of recording with samples of two Portuguese spoken dialects in Amazonia Region: the Amazon Portuguese (AM) and the Afro-Brazilian Portuguese (ABP). As a baseline, samples of the Standard Brazilian Portuguese (PB) were also used. The corpus inventory, the digitalization of the speech data and the phonetic segmentation using the PRAAT software preceded the construction of the oral database. Using language HTML and other resources of computer science a documentation of the database in a Internet site becomes possible.

1. INTRODUÇÃO

O objetivo principal do presente estudo é o de demonstrar a importância de utilização de dados espontâneos nos estudos acústico-experimentais, assim como o da conciliação de *corpus* sociolingüísticos e da metodologia fonético-experimental.

A proposta de trabalho aqui descrita nasceu das dificuldades encontradas por Cruz (2000b) na exploração da totalidade de seu *corpus* e que impossibilitaram um aprofundamento nos processos estudados. Trata-se, portanto, da atualização de um banco de dados orais concebido e já executado desde Cruz (2000b). A organização e disponibilidade desse *corpus* num banco de dados destaca-se na sua originalidade por ser uma base de dados orais constituída de fala espontânea, útil para estudos de variação lingüística do português.

A necessidade de documentar e disponibilizar eletronicamente dados de fala espontânea representativos do português brasileiro decorre da quase total inexistência de *corpora* organizados em forma de banco de dados orais e disponíveis eletronicamente a fim de servir de suporte para toda e qualquer investigação lingüística, assim como para aplicação tecnológica envolvendo engenharia de fala.

O crescente interesse dos estudos fonéticos em relação à fala espontânea não é suficiente para uma aplicação imediata e direta da metodologia da fonética experimental a dados de fala espontânea. A impossibilidade maior está na forte tradição de utilização da fala lida ou paradigmaticamente gerada nas experimentações em fonética refletida, por exemplo, na alta sensibilidade dos equipamentos utilizados nesse tipo de estudo.

No sentido de esclarecer melhor tais dificuldades, uma descrição detalhada das mesmas é feita ao longo do presente estudo, assim como dos procedimentos empreendidos na tentativa de facilitar o

acesso automático a dados do português espontâneo por estudos na área de fonética experimental e engenharia de fala, através da organização de um *corpus* em banco de dados orais, contendo informações qualitativas e quantitativas dos sinais sonoros armazenados. Qual então é o *corpus* em questão?

2. CORPUS

O referido *corpus* compreende mais de 30 horas de gravação com amostras de três variedades lingüísticas do português: (i) o português regional paraense (AM), coletado no período de 1989–1990 (TRINDADE); (ii) o português afro-brasileiro (ABP), variedade alvo, falada pelas comunidades quilombolas do Pará, coletado entre 1992-1996 (CRUZ, 2000b) e; (iii) o português brasileiro padrão (PB) emprestado de Oliveira (2000), com amostra da variedade lingüística em grandes centros urbanos e por falantes de alto nível de escolaridade.

Dois aspectos os aproximam:

- a) a língua: todos contêm amostras do português brasileiro;
- b) a metodologia seguida para sua formação: os três *corpora* foram formados prioritariamente para servir de suporte empírico a investigações sociolingüísticas.

Inicialmente, todos os sinais sonoros armazenados nesse banco de dados orais foram coletados em fitas cassete em trabalho de campo. Das quais foram retiradas as informações necessárias acerca do *corpus*. Antes da digitalização dos sinais sonoros, Cruz (2000b) procedeu a um inventário das cassetes de áudio. Esse inventário permitiu identificar quais seriam as gravações mais apropriadas para os estudos fonéticos. A escuta permitiu chegar a uma classificação das fitas em três categorias indicadas através de cores. Essa classificação repousa sobre a qualidade sonora das gravações que varia de uma tomada de gravação¹ a outra.

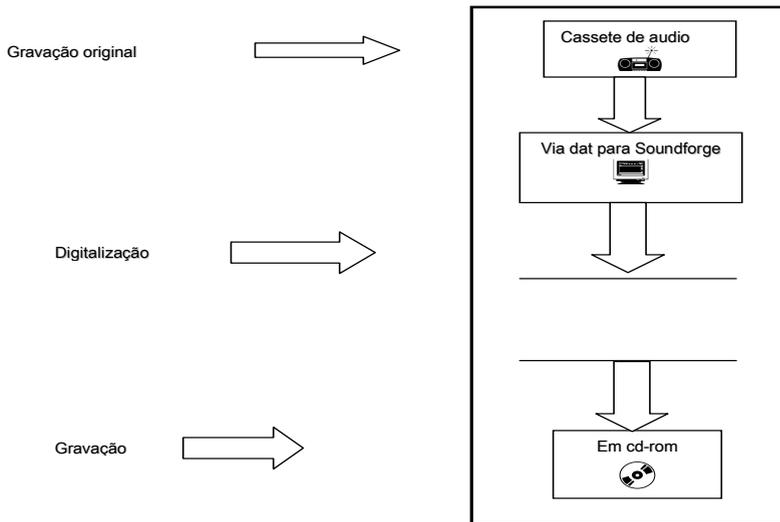
Esta mesma classificação é retomada em um arquivo específico do banco de dados orais aqui descrito. As tomadas de gravação

foram então identificadas como: <<azul>> (tomadas de gravação de excelente qualidade sonora); <<amarelo>> (tomadas de gravação cuja utilização para análises acústicas é duvidosa) e <<vermelho>> (tomadas de gravação cuja situação de fala não faz parte daquelas consideradas apropriadas para um estudo sobre fala espontânea ou tomadas de gravação inutilizáveis para um estudo acústico). Além de precisar sobre a qualidade das gravações, o inventário feito por Cruz (2000b) também procedeu a um levantamento acurado sobre faixa etária, sexo e procedência dos locutores, situação de fala, local das gravações e variedade lingüística. Todas essas informações figuram no código atribuído a cada sinal sonoro no momento da digitalização.

3. DIGITALIZAÇÃO

Uma exigência para facilitar o tratamento experimental foi a transferência do *corpus* para uma base fixa. Para esta fase, também o inventário preliminar do *corpus* foi fundamental. Para a transferência do *corpus* para CDROM, Cruz (2000b) decidiu fazer a transferência de todas as fitas cassetes classificadas como AZUL e AMARELO, bem como as gravações classificadas como VERMELHO, que não apresentavam problemas de rotação da gravação.

A digitalização das gravações em sinais de áudio foi feita no programa SOUNDFORGE via um gravador DAT. Os parâmetros utilizados por Cruz (2000b) foram: 44100 Hz, 16 bit, Mono. A saturação do sinal foi controlada por um monitor de maneira que o Tbs áudio remixador permitisse verificar se a digitalização estava correta. O esquema 1 explica a maneira como foi feita a digitalização.



Esquema 1 – Fases do processo de digitalização das gravações
(extraído e traduzido de CRUZ 2000b: 183)

O resultado desta primeira fase foi a gravação de 38 CD-ROMS áudio que totalizaram 39 horas, 19 minutos e 55 segundos.

A fase de digitalização foi extremamente importante para a organização dos dados. A unidade “tomada de gravação” estabelecida como informação pertinente para o *corpus* foi retomada nesta fase. A noção de tomada de gravação aparece pela primeira vez no momento de preparar o inventário do *corpus*: o critério utilizado para a delimitação destas tomadas de gravação é simplesmente o fato de ligar e desligar o gravador. Por esta razão, «tomada de gravação» foi empregada com o sentido de «evento sonoro gravado sem interrupção». Elas são a unidade base que compõe os sinais sonoros armazenados

O programa SOUNDFORGE permitiu reunir duas ou mais tomadas de gravação que pertenciam a uma mesma situação de fala e que, por razões técnicas (limitação física do tempo da fita cassete por exemplo), se encontravam em tomadas de gravação separadas. No caso de uma mesma fita cassete conter várias situações de fala diferentes, as quais nem apresentavam relação direta, foi possível a

separação das mesmas; assim como foi possível também reunir em um mesmo sinal, por exemplo, todas as tomadas de gravação que pertenciam à fala lida. Estas foram reunidas em um mesmo sinal e separadas por um intervalo de silêncio de 600 milissegundos. Foi igualmente possível eliminar barulhos indesejáveis para análise acústica.

Para não se perder a vinculação com as gravações originais, cada sinal gravado contém as seguintes informações: um número de ordem, a data da gravação e a duração total do sinal de áudio que ele contém, além de detalhes de cada arquivo digitalizado por SOUNDFORGE (assinalando data da digitalização, as tomadas de gravação originais, e a duração do sinal .wav obtido). As outras informações estão disponíveis no próprio código atribuído a cada sinal .wav, o qual indica: a situação de fala, sexo do locutor, lugar e qualidade da gravação.

A digitalização com SOUNDFORGE e a transferência para CDROM permitiu organizar melhor os sinais sonoros, entretanto, estes ainda não estavam prontos para o tratamento com os programas de análise acústica experimental. Era necessário que os CDs áudio fossem transformados em CDs de dados e para essa tarefa foi utilizado o programa COOLEDT. Cruz (2000b) fez a redigitalização dos sinais sonoros de acordo com os seguintes parâmetros: frequência de 16kHz, 16 bits e sinal mono. Desta forma, o COOLEDT fornecia um sinal .wav pronto para ser trabalhado com os programas de tratamento acústico. Nem todos os sinais áudio de 44100 hz foram redigitalizados em 16 kHz. A autora selecionou apenas aqueles contendo mais amostras de fala espontânea, assim como os sinais que apresentavam melhor qualidade de gravação. Por esta razão, ela escolheu, em um primeiro momento, 61 sinais de áudio a serem redigitalizados.

Houve também a redigitalização de 7 sinais não pertencentes a fala espontânea (1 de celebração religiosa e 6 de fala lida) com o objetivo de estabelecer comparações futuras com outras situações de fala. No total, 5 CDs de dados foram gravados.

Todo este trabalho de transferência dos dados para CDROM permitiu trabalhar os sinais tais quais se apresentavam nos dados em

tomadas de gravação, evitando, assim, a digitalização direta nos programas de análise acústica, pois mesmo o PRAAT, que permite analisar sinais de mais de 3 horas de duração, restringe a duração a um minuto por aquisição de sinal. Mesma situação Cruz (2000b) enfrentou com o programa MES, que permitia apenas um minuto de aquisição direta. Desta maneira, no caso do PRAAT foi possível fazer a leitura do sinal diretamente do CDROM.

Se um *corpus*, que não foi concebido inicialmente para sustentar um estudo experimental, mas com a intenção de ser representativo da variedade lingüística estudada, permite fazer análises acústicas, é possível que esta proporção aumente ainda mais em um *corpus* formado com a intenção, desde o início, de se obter o espontâneo para estudos experimentais. O *corpus* do PB que foi emprestado é um exemplo disso, pois o autor formou um *corpus* representativo do PB com uma alta fidelidade de gravação (OLIVEIRA, 2000).

A etapa de digitalização foi um trabalho essencial para a construção desse banco de dados, já que permitiu obter sinais sonoros prontos para serem tratados em programas de análise acústica. O intuito não é apenas manter um acervo documentado do português brasileiro, interessa também tornar o acesso a esses dados facilitado. E a forma mais viável encontrada foi a organização do *corpus* em um banco de dados orais, o qual será descrito a seguir.

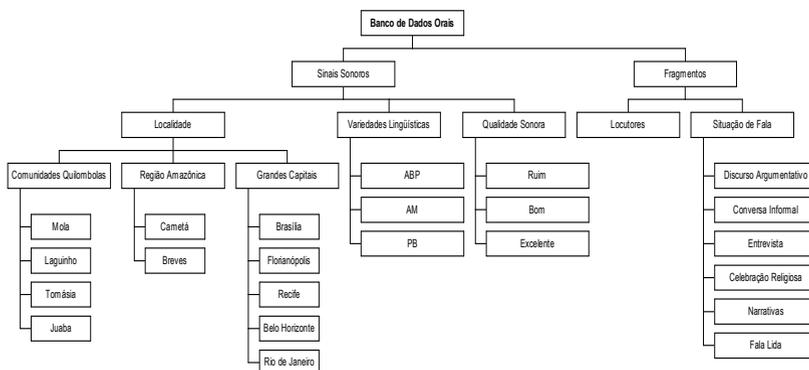
4. DESCRIÇÃO DO BANCO DE DADOS

Como se tratava de um *corpus* coletado em trabalho de campo, antes da concepção e posterior execução do banco de dados, como já mencionado acima, foi feito um inventário do material lingüístico presente no mesmo (CRUZ, 2000b). A informação levantada: número de locutores, a diversidade de situação de fala, tempo, localidades e datas das gravações foram importantes para a construção do banco de dados, que contém informações tanto qualitativas, quanto quantitativas sobre as gravações. Vejamos como estas informações estão disponíveis no banco de dados concebido para permitir o manuseio automático desse material lingüístico.

O banco de dados é composto de sete arquivos interligados: (i) o das localidades, contendo informações sobre os povoados e as cidades de procedência dos locutores; (ii) o dos locutores; (iii) o dos sinais sonoros, contendo informações diretas sobre as tomadas de gravações e; (iv) um arquivo chamado fragmento. Este último arquivo contém as informações qualitativas (características do locutor, situação de fala, e também transcrição fonética e lexical do sinal sonoro) e informações quantitativas (medidas físicas sobre gravações, as segmentações realizadas, o sinal sonoro); (v) o das variedades lingüísticas; (vi) o da qualidade dos sinais sonoros; (vii) o da situação de fala.

Por se tratar de um banco de dados do tipo relacional, inicialmente foi proposto a utilização do programa FILEMAKER PRO.4 na execução do banco de dados orais (CRUZ, 2000b; CRUZ, HIRST & BEL). Todavia, constatando uma incompatibilidade entre o FILEMAKER PRO.4 e o programa PRAAT, utilizado para a segmentação dos sinais sonoros em palavras e fonemas, preferiu-se tentar a edição de um *site* na *Internet*, usando linguagem HTML e outros recursos de animação sonora.

O esquema 2 mostra como o banco de dados orais está estruturado hierarquicamente.



Esquema 2 – Estrutura hierárquica do banco de dados

Como podemos perceber, este banco de dados é composto por arquivos interligados que apresentam informações tanto qualitativas, quanto quantitativas dos sinais sonoros. Além da página inicial, o banco de dados orais é composto dos seguintes arquivos:

4.1. Sinais Sonoros

Este arquivo contém informações diretas sobre os sinais sonoros, que correspondem à identificação do locutor, à duração do sinal sonoro, à situação de fala, à data, ao lugar e à qualidade de gravação. Assim como disponibiliza uma segmentação em turnos de fala de cada sinal sonoro e possibilita uma comprovação auditiva desta mesma segmentação em um arquivo .mp3. Estes sinais sonoros correspondem à unidade tomada de gravação descrita no item 2.

O gráfico 1 mostra a porcentagem de sinais sonoros armazenados já transcritos em turnos de fala.

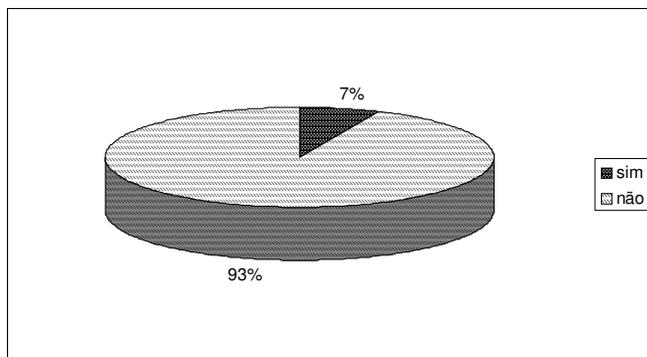


Gráfico 1 – porcentagem de sinais sonoros armazenados já transcritos em turno de fala (total de 61 sinais sonoros).

Para figurar no banco de dados, os sinais sonoros foram codificados segundo um padrão de notação cuja convenção está descrita no quadro 1 e na figura 1 abaixo:

Variedade lingüística	Local	N.º	Tomada	Gravação	Situação de fala	Qualidade da gravação
A: Afro Am: Amazonia B: Brasileiro P: Português	M: Mola J: Juaba T: Tomásia L: Laguinho C: Cametá Cv: Custavê A: Ajó I: Itapocu BR: Brasília CR: Curitiba RJ: Rio de Janeiro BH: Belo Horizonte FL: Florianópolis				E: entrevista CI: conversa informal N: narrativa CR: Celebração religiosa TA: discurso argumentativo L: fala lida X: não identificado	E: excelente (= azul) B: bom (= amarelo) R: ruim (= vermelho)

Quadro 1 – todas as informações utilizadas na notação dos sinais sonoros (emprestado e traduzido de CRUZ, 2000b: 267)

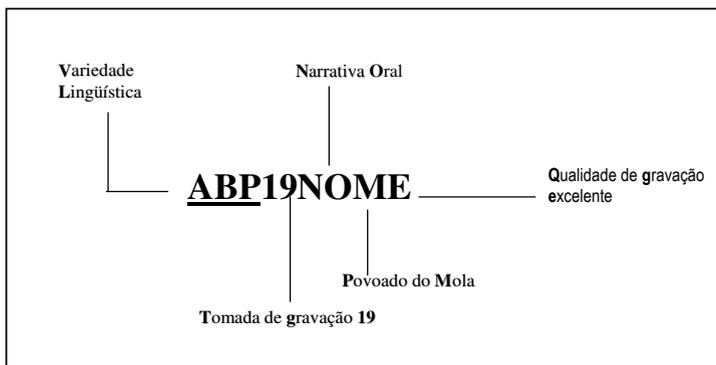


Figura 1 – tomadas de gravação foram codificadas segundo as seguintes convenções (emprestado e traduzido de CRUZ, 2000b: 291): *ABP19NOME* = Variedade lingüística (*ABP*); tomada de gravação (*19*); situação de fala (*narrativa oral*); localidade (*povoado do Mola*); qualidade de gravação (*Excelente - E*).

O arquivo de sinais sonoros contém as seguintes informações:

4.1.1 Localidades

Este arquivo contém prioritariamente informações sócio-econômicas e histórico-culturais das comunidades quilombolas de Juaba, Mola, Tomásia e Laginho, localizadas no município de Cametá (PA). No banco de dados há referência tanto à localidade de origem dos locutores, quanto às localidades onde foram feitas as gravações.

4.1.2 Variedades Lingüísticas

Este arquivo contém informações sobre as três variedades do português falado no Brasil com amostras no banco de dados.

- a) **AM** – português regional da Amazônia [TRINDADE]. Foi também o *corpus* utilizado para as primeiras experimentações com fala espontânea;
- b) **PB** – português brasileiro padrão. *Corpus* complementar emprestado de Oliveira [7] para efeito de comparação;
- c) **ABP** – português afro-brasileiro como definido por Cruz (2000b).

O gráfico 2 apresenta como as variedades estão distribuídas no banco de dados em número e duração de sinais sonoros.

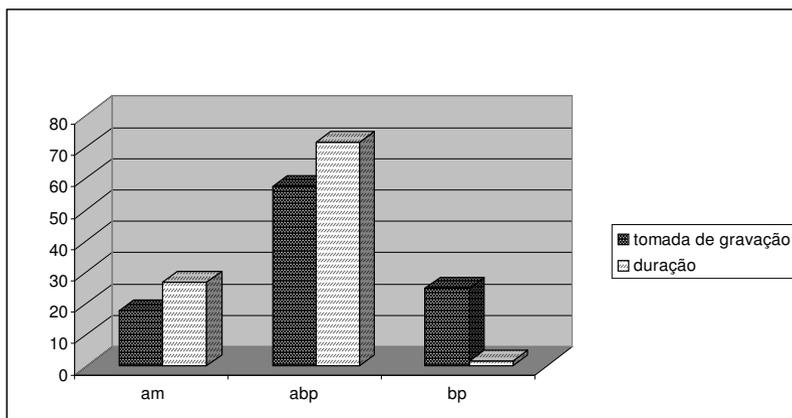


Gráfico 2 – visão do *corpus* em variedades lingüísticas, tomada e duração de gravação (total de 61 sinais sonoros e duração total de 69.760 segundos) (emprestado e traduzido de CRUZ 2000b:185)

Os *corpora* foram formados em momentos distintos. O português regional da Amazônia foi o primeiro a ser formado, sua coleta de dados se deu entre setembro de 1989 a fevereiro de 1990 [8]. O português afro-brasileiro foi coletado durante três trabalhos de campo: (i) novembro a dezembro de 1992; (ii) outubro de 1993 a janeiro de 1994 e (iv) setembro a dezembro de 1994 (CRUZ, 2000b). Os dados do português brasileiro padrão foram gravados em 1998 com falantes do português residentes em Vancouver no Canadá (OLIVEIRA, 2000), todos de classe média alta e com nível de escolaridade alto. Foram dados coletados com o objetivo de se estudar a fala espontânea do ponto de vista experimental e numa visão laboviana de trabalho de campo.

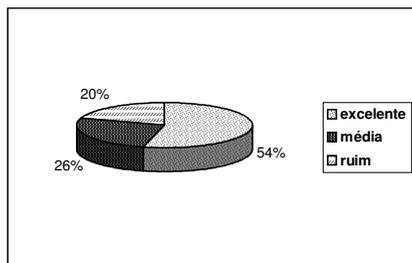
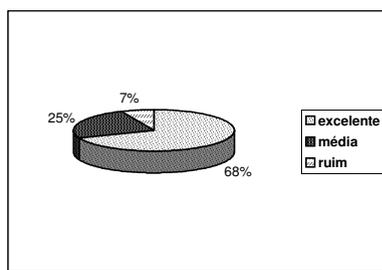
Este arquivo apresenta informações sociolingüísticas e lingüísticas sobre as referidas variedades, assim como os resultados de Cruz (2000b) e Oliveira (2000).

4.1.3. Qualidade Sonora

Descreve-se como as tomadas de gravações foram transformadas em sinais sonoros, tomando o critério da qualidade das gravações, e indica quais são os sinais sonoros de *excelente qualidade de gravação* e quais *não são aconselháveis* para um estudo acústico-experimental.

O gráfico 3 mostra o *corpus* original ainda em cassete, e o gráfico 4 mostra o *corpus* final digitalizado. Ressaltando que, no gráfico 3, os 20% de gravação considerados ruim indicam tanto problemas de ordem técnica (rotação irregular, muito barulho), quanto dados não representativos de fala espontânea. Enquanto o gráfico 4 indica que 7% das gravações etiquetadas como VERMELHO são amostras de fala não-espontânea.

Apesar de se tratar de um *corpus* não concebido especificamente para sustentar um estudo fonético experimental, obtiveram-se tomadas de gravação passíveis de serem utilizadas nesse tipo de estudo. Logo, essa proporção deve aumentar em coleta de dados, visando obter tanto fala espontânea, quanto gravações com alta fidelidade.

Gráfico 3 - *corpus* original ainda em casseteGráfico 4 - *corpus* final digitalizado

4.2. Fragmento

Este arquivo é o mais importante para quem está realizando estudos acústicos, pois contém informações de acesso aos segmentos sonoros do *corpus* armazenado. Aqui um fragmento sonoro corresponde a um turno de fala, um só enunciado de um único locutor. Os resultados das análises acústicas, como: duração, intensidade e frequência, estão estocadas neste arquivo. Em termos de tamanho, é o arquivo mais volumoso, pois ele contém um grande número de fichas. As fichas contêm as informações que permitem o acesso ao sinal acústico de formato .mp3, e a sua segmentação em palavras e fonemas. No que diz respeito à segmentação feita pelo alinhador do MBROLIGN, estas foram convertidas para o formato TEXGRID, aceito por PRAAT, através de *script* PERL e sofreram revisão manual.

Além do sinal sonoro em si, nós podemos acessar as informações contextuais mais precisas sobre os sinais sonoros tratados de forma automática.

Os fragmentos sonoros são transcritos lexical e foneticamente, com os quais é possível realizar as análises prosódicas de ritmo e entoação, pois sua unidade de base compreende a voz de único locutor. Eles estão sendo transcritos com PRAAT para a parte segmental e a codificação MOMEL para a normalização da curva de F0 (CRUZ, 2000b). A transcrição segmental é feita com o alfabeto SAMPA. Há 80 fragmentos sonoros segmentados até o momento.

Este arquivo contém os seguinte sub-arquivos:

4.2.1. Locutores

Este arquivo apresenta informações sobre sexo, faixa etária, origem, escolaridade e variedade lingüística dos locutores.

Os gráficos abaixo mostram informações dos locutores quanto ao sexo e à faixa etária. O gráfico 6 especifica que, no *corpus* há um número maior de vozes masculinas. E o gráfico 7 mostra que há uma maior representatividade de locutores na faixa de 30-70 anos de idade.

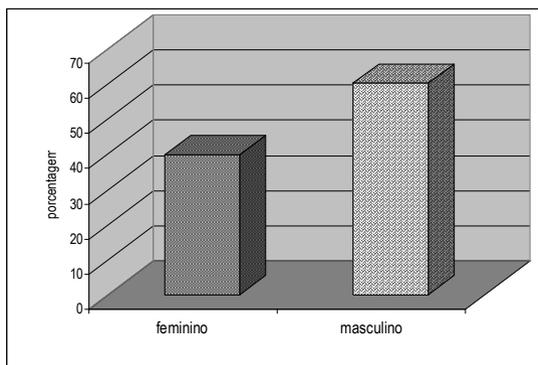


Gráfico 5 - informações dos locutores quanto ao sexo

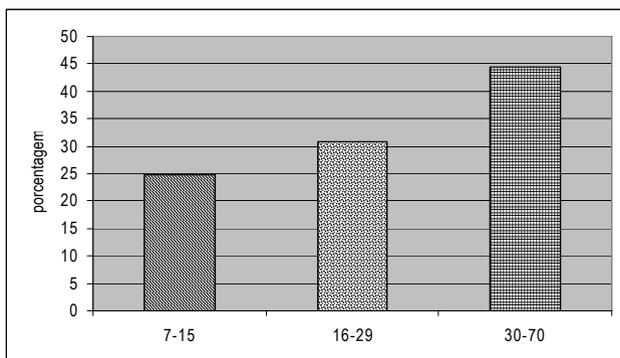


Gráfico 6 - distribuição dos locutores por faixa etária

4.2.2 Situação de Fala

Apresenta a distribuição dos 61 sinais sonoros em situação de fala, como: discurso argumentativo, conversa informal, entrevista, celebração religiosa, narrativas e fala lida.

Foram escolhidos, em um primeiro momento, 61 sinais de áudio, distribuídos da seguinte forma como expresso no gráfico 7.

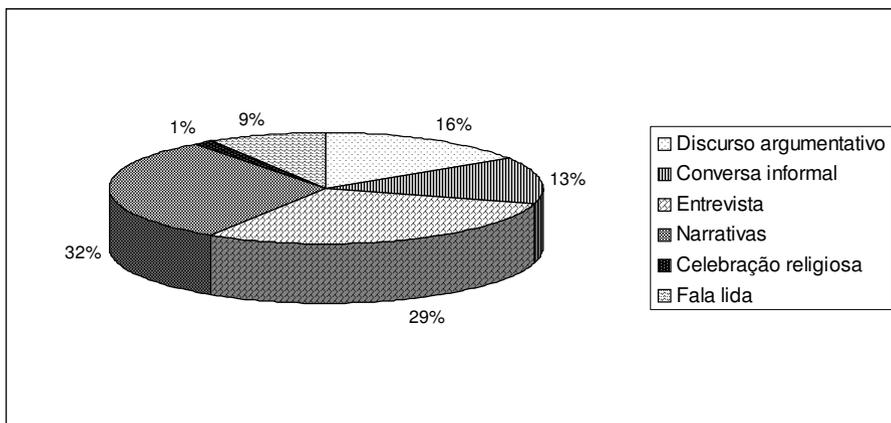


Gráfico 7 - composição dos Cds de dados em situação de fala

5. COMO INTERROGAR O BANCO DE DADOS?

Não há uma linguagem de interrogação específica para utilizar este banco de dados, pois os operadores lógicos necessários a uma busca estão disponíveis diretamente na interface gráfica do *site Internet*.

Este banco de dados apresenta informações diretas sobre os sinais sonoros, que correspondem a: identificação do locutor, data e lugar de gravação, qualidade da gravação, situação de fala, e duração do sinal. Encontra-se disponível, ainda, uma segmentação em turnos de fala de cada sinal sonoro, contendo as medidas de duração e frequência referentes a cada segmento em um arquivo ASCII. Também é possível uma comprovação auditiva desta mesma segmentação em um arquivo mp3.

6. DESCRIÇÃO DO SITE

O usuário do banco de dados poderá acessá-lo através do *site* que o hospeda e que possui, como “porta de entrada” e página principal, uma *Home Page*, contendo a página principal do *site*, e como tal funciona como portal para todas as outras páginas, permitindo assim que estas estejam interligadas entre si. Em sua totalidade o *site* apresenta 23 documentos em HTML² que apresentam descrições detalhadas sobre o trabalho desenvolvido no projeto, sobre os integrantes da equipe de trabalho, assim como dos colaboradores e consultores científicos. Também apresenta uma página com formulário que permite ao cliente (usuário) interagir com o servidor, preenchendo campos, clicando em botões e passando informações, o que possibilita ao visitante entrar em contato com a coordenação do projeto e emitir suas opiniões. E tem como documento principal uma página intitulada Banco de Dados Orais, que funciona como um *link* que interliga a *Home Page* ao banco de dados orais, permitindo, assim, uma perfeita interação entre *site* e os arquivos do banco de dados.

Com a construção do banco de dados orais e sua implementação na rede *Internet* prevemos ainda a *interface* entre a sede do banco de dados em HTML³ que dispõe de informações qualitativas, e os programas de análise acústica e estatística (PRAAT e MES), nos quais está sendo feito o trabalho de segmentação dos sinais.

A *interface* entre o *site* e esses programas permite ao usuário conhecer o método de análise acústica anterior a implementação dos *corpus* orais num hospedeiro da *Internet*. A figura 2 mostra a interligação entre o *site* e os programas de análise acústica.

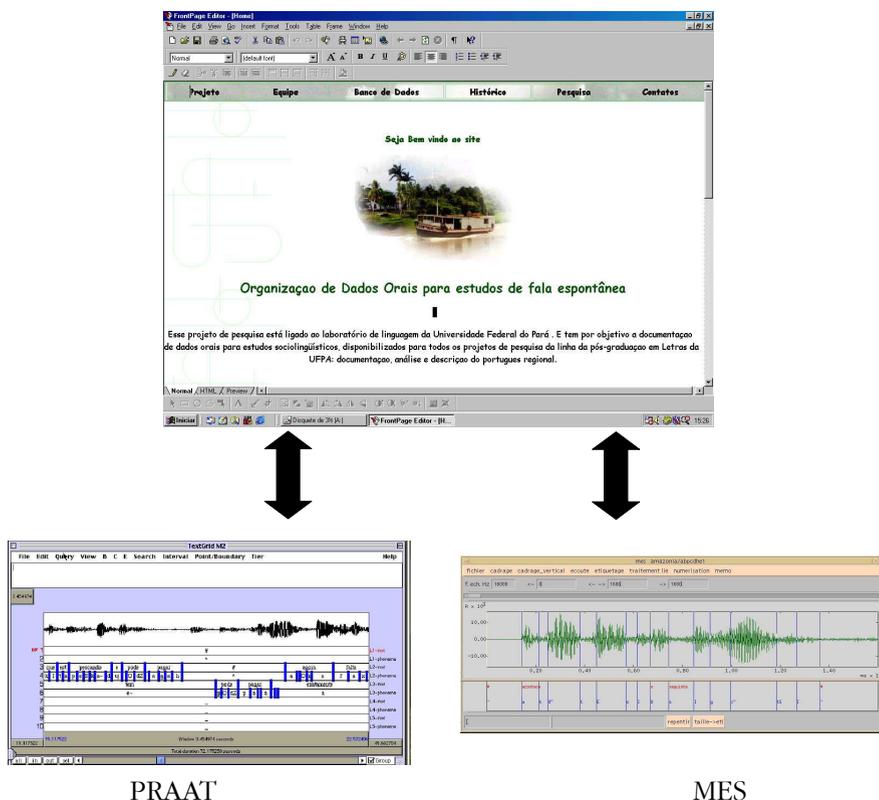


Figura 2 - interligação entre o *site* e o banco de dados e os programas de análise acústica.

Na fase atual, a aplicação e manutenção desse *site* na rede possibilita a divulgação desse trabalho destinado a fornecer informações lingüísticas para os estudiosos da linguagem e áreas afins. Seu caráter inovador se destaca por aplicar a *web* um banco de dados orais contendo fala espontânea, o que até então não poderia ser encontrado na rede, já que a grande maioria dos *sites* que tratam de *corpora* orais na *Internet* não conceberam tal empreendimento.

7. QUESTÕES TEÓRICAS

Diferentes escolas lingüísticas atestam uma importância ao status do *corpus* nos estudos lingüísticos (estruturalistas, gerativistas e sociolingüísticas). Por volta dos anos 50, somente Chomsky colocou em discussão a importância do uso de *corpora* como fonte primária de informação para as investigações lingüísticas. Todas as outras escolas atribuíram uma importância capital aos dados, apesar do fato de que elas não partilhavam da mesma concepção de *corpus* lingüístico em si (McENERY & WILSON, 1997).

A fonética, enquanto ramo da lingüística, apresenta este mesmo quadro. Sociolingüistas e foneticistas foram os primeiros a estudar a fala como objeto de estudos lingüísticos, particularmente quando eles se voltam para a fala espontânea. Assim, nós poderíamos estabelecer uma relação entre o vernáculo da sociolingüística e a fala espontânea para a fonética.

Entretanto, os estudos fonético-experimentais ainda se utilizam de fala lida de laboratório ou dados fonoestilísticos ou mesmo ainda dados produzidos em conversação controlada em laboratório em suas investigações. A causa maior é em decorrência da sensibilidade dos programas de análise que não são capazes de distinguir entre fala e os outros tipos de sons. Portanto, a qualidade das gravações compromete a naturalidade da fala. Eis a razão da utilização de gravações de alta fidelidade sem barulho anexo.

A acuidade fornecida pelas análises experimentais sobre os estudos dos sons da fala é inegável. Constitui-se um forte instrumento

sofisticado de precisão disponível para a pesquisa, quando nós precisamos tirar dúvidas que o ouvido humano não é capaz de solucionar. Este tipo de informação torna claro e dá conta de questões extralingüísticas indecifráveis por teorias fonológicas.

O problema maior do uso de dados espontâneos em trabalho experimental é a exigência de *corpus* de tamanho considerável, pois em sua totalidade são *corpus* com um controle bem menor do que os de fala artificial.

Entretanto, a fonética experimental deve ser contaminada pela concepção de língua da sociolingüística, e trabalhar muito mais com dados de língua em uso, em comunicação, em interação (LABOV, 1976).

A importância de banco de dados em tecnologia de fala é sentida em todos os níveis: (i) a disponibilidade de fonte de registro de fala devidamente documentado pode ser de grande utilidade para estudos diacrônicos futuros; (ii) a síntese de fala e o reconhecimento de voz tendem a utilizar muito mais sistemas <<data driven>>; (iii) o rumo epistemológico dos estudos sobre a linguagem, que priorizam muito mais conhecer a interrelação entre os níveis do que os níveis em si, como dominou por muito tempo na lingüística e; (iv) o número cada vez maior de estudos interdisciplinares sobre a linguagem humana.

8. CONCLUSÃO

A organização e disponibilidade desse tipo de *corpus* é de extrema relevância para a comunidade científica. Poucos são aqueles projetos cujo *corpus* foi organizado em forma de base de dados orais e encontra-se disponível eletronicamente. E como se trata de um banco de dados orais destinado a estudos de variação lingüística, nada melhor do que documentá-lo devidamente e disponibilizá-lo para todos os projetos de pesquisa envolvendo o português falado. E já que, atualmente, são feitos investimentos pesados em *Internet*, acredita-se ser de suma importância a iniciativa de disponibilizar esses bancos de dados orais num *site* da rede *web*.

NOTAS

- ¹ Unidade sonora ininterrupta.
- ² <http://www.fbs.aust.com/aardvark.html>. editor html com várias funções;
<http://www.infoflex.com.au/flexed.htm>. editor html;
<http://www.sede.com>. sede *Internet* – serviços de hospedagem.
- ³ <http://www.fbs.aust.com/aardvark.html>. editor html com várias funções;
<http://www.infoflex.com.au/flexed.htm>. editor html;
<http://www.sede.com>. sede *Internet* – serviços de hospedagem.

REFERÊNCIAS BIBLIOGRÁFICAS

CRUZ, R. Setting up spontaneous speech corpora. Comunicação oral apresentada no *Workshop Méthodes et Formalismes pour la Linguistique de Corpus*, realizado em Aix-en-Provence, Faculté des Lettres, Salle des Professeurs, no período de 12-13 octobre 2000a.

CRUZ, R. *Aspects phonologiques et acoustiques du portugais parlé par des communautés noires de l'amazonie (Brésil)*. 2000b. Thèse (Doctorat) - Université de Provence.

CRUZ, R.; HIRST, D. Mise en oeuvre d'un corpus spontané. Trabalho apresentado em forma de painel no *RJC99 (Rencontre des Jeunes Chercheurs en Parole)*, Avignon (France), IUP Informatique d'Avignon, 18-19 de novembro, 1999.

CRUZ, R.; HIRST, D.; BEL, B. Mise en oeuvre d'un corpus spontané. *Revue Parole*. DUEZ, Danielle (Ed.). volume especial sur la parole spontanée.

LABOV, W. *Sociolinguistic*. Traduction d'Alain Kihm. Les Éditions de Minuit: Paris, 1976.

McENERY, T.; WILSON, A. *Corpus linguistic*. Edimburg University Press: Edinburg, 1997.

OLIVEIRA, M. *Prosodic Features in spontaneous narratives*. 2000. 241f. Dissertation (PhD) - Simon Frase University, 2000.

TRINDADE, R. *O som da fala dos pescadores de Cameté*. Memóire de Master du Département de Linguistic de l'Université Fédérale du Santa Catarina (Brésil).

Teoria Fonológica

O caminho para a realização de / l / e / r / intervocálicos na aquisição com atraso: a substituição por [j]

José Olímpio de Magalhães
Universidade Federal de Minas Gerais

Maíra Martins
Universidade Federal de Minas Gerais / UNIPAC

Abstract

The main purpose of this paper is to describe and analyse the linguistic behavior in delayed acquisition by four children, aged from 4:1 to 7:4, replacing intervocalic /l/ and /r/ by [j], in order to explain the path they take to realize intervocalic /l/ and /r/. The most common strategies in producing intervocalic /l/ and /r/ by these children to reach the adult model have been observed and, out of this, a hypothesis about the taken path has been proposed: first they don't realize intervocalic /l/ and /r/; second, they change intervocalic /l/ and /r/ by [j]; third, they realize intervocalic /l/; fourth, they replace /r/ by [l]; finally, they realize intervocalic /r/. The analysis of the results confirmed that children follow this path but not in a systematic way, because the substitution stage of [l] by /r/ is optional.

1. INTRODUÇÃO

Um dos objetivos deste estudo é descrever e analisar as estratégias de produções¹ de /l/ e /r/ intervocálicos na fala de quatro crianças, na faixa etária dos 4:1 aos 7:4 anos de idade, que substituem /l/ e /r/ intervocálicos por [j], em termos das não-realizações,² substituições e realizações, investigando os fatores envolvidos e as implicações da busca do alvo adulto. Dessa forma, pretende-se estabelecer um perfil de comportamento lingüístico na aquisição³ com atraso de /l/ e /r/ intervocálicos, verificando a influência da tonicidade, das vogais antecedentes e seguintes, da posição na palavra e do número de sílabas em cada etapa do caminho rumo à realização de /l/ e /r/ intervocálicos.

Outro objetivo é a descrição, análise e explicação de cada etapa do caminho percorrido, à luz da geometria de traços de Clements & Hume (1995) e dos modelos implicacionais de complexidade de traços, seguindo proposta de Mota (1996) e Rangel (1998), que permitem estabelecer a hierarquia dos traços necessária à aquisição dos segmentos-alvo.⁴

Na busca de seus objetivos, este estudo partiu, pois, das seguintes hipóteses:

- a) O /l/ é realizado antes de /r/.
- b) /l/ e /r/ podem ser não-realizados, substituídos ou realizados.
- c) Quando encontramos /l/ substituído por [j], encontramos também o /r/ substituído por [j].
- d) Se o informante já realiza o /l/, ele não irá mais substituir o /r/ por [j], e sim por [l].
- e) A posição pós-tônica é a mais atingida pela substituição de /l/ e /r/ por [j].

- f) Tanto as vogais antecedentes quanto as seguintes exercem influências nas estratégias de produção do /l/ e /r/.
- g) Em posição não-inicial na palavra há mais substituições por [j].
- h) Quanto maior a palavra, maior o número de não-realizações de /l/ e /r/.

O trabalho segue a seguinte ordem:

Após essa INTRODUÇÃO, tem-se, na 2^a. parte, a seguir, a metodologia utilizada: dos informantes; da coleta dos dados; da transcrição dos dados; do material coletado.

Na terceira parte, são apresentados os resultados e a análise dos dados obtidos na pesquisa.

Na última parte, à guisa de conclusão, são retomadas as hipóteses iniciais acerca do caminho proposto para a realização de /l/ e /r/ intervocálicos.

2. METODOLOGIA

A idéia central deste estudo é propor um caminho para a realização de /l/ e /r/ intervocálicos na aquisição com atraso, percorrido por crianças que substituem /l/ e /r/ intervocálicos por [j]. Diante disso, partimos da hipótese de que o caminho para a realização do /l/ intervocálico, nestes casos, se faz da seguinte forma, consecutivamente:

- *Etapa 1:* /l/ → 0: onde o /l/ não é realizado. Ex: “bola” produzida como [b^hbə]
- *Etapa 2:* /l/ → [j]: onde /l/ é substituído por [j]. Ex: “bola” produzida como [b^hbjə]
- *Etapa 3:* /l/ → [l]: onde /l/ é realizado como [l]. Ex: “bola” produzida como [b^hlə]

E para o /r/ intervocálico, da seguinte forma, também consecutivamente:

- *Etapa 1:* /r/ → 0: onde /r/ não é realizado. Ex: “pêra” produzida como [p^hə]
- *Etapa 2:* /r/ → [j]: onde /r/ é substituído por [j]. Ex: “pêra” produzida como [p^hejə]
- *Etapa 3:* /r/ → [r]: onde /r/ é substituído por [r]. Ex: “pêra” produzida como [p^helə]
- *Etapa 4:* /r/ → [r]: onde /R/ é realizado como [r]. Ex: “pêra” produzida como [p^herə]

Mostramos os caminhos propostos para a realização do /l/ intervocálico em 3 etapas, e, do /r/ intervocálico, em 4 etapas, para efeito de análise isolada; entretanto, a idéia principal deste estudo é propor somente um caminho para a realização destes dois segmentos, por considerarmos que a realização de /r/ depende da realização de /l/. Este caminho único contém 5 etapas, descritas a seguir:

- *Etapa 1:* /l/ e /r/ → 0: onde o /l/ e /r/ não são realizados.
- *Etapa 2:* /l/ e /r/ → [lj]: onde /l/ e /r/ são substituídos por [lj].
- *Etapa 3:* /l/ → [l]: onde /l/ é realizado.
- *Etapa 4:* /r/ → [r]: onde /r/ é substituído por [r].
- *Etapa 5:* /r/ → [r]: onde /r/ é realizado.

Os informantes desta pesquisa, num primeiro momento, não conseguem realizar foneticamente /l/ e /r/ (etapa1), embora consigam discriminar as palavras “sala” e “Sara”. Além disso, estão cientes de que existe “algo” na palavra que eles não conseguem produzir ou produzem incorretamente. Tal fato indica que estes informantes possuem os segmentos-alvo fonologicamente.

O próximo passo, na tentativa de se atingirem os segmentos-alvo, é a substituição por um segmento já adquirido e, portanto, mais fácil e que se assemelhe, de alguma forma, com os segmentos-alvo. Entre as vogais, [i] é a que mais se assemelha, do ponto de vista articulatorio, a /l/ e /r/ por se tratar de uma vogal alta e anterior; porém, por definição, só pode ocupar núcleo de sílaba. Representamos a substituição de /l/ e /r/ pelo vocóide [j] (etapa 2) que compartilha das mesmas propriedades que a vogal [i], com a diferença de que pode ocupar a posição de *onset* na sílaba.

O último passo é que, uma vez adquirido o /l/ (etapa 3), a criança passa a não substituir mais o /r/ por [lj], e sim por [l] (etapa 4). A criança consegue perceber os dois segmentos como pertencentes à mesma classe, das líquidas. Isto ocorre até que a criança consiga realizar foneticamente o [r] (etapa 5).

Nem todas as crianças deste estudo passarão, necessariamente, por todas estas etapas, o que implica que algumas etapas possam

co-ocorrer. A criança, por exemplo, pode não realizar o /l/ em alguns contextos, realizá-lo ou substituí-lo por [j] em outros, justamente porque ela se encontra em fase de aquisição e está “experimentando o caminho”. A estratégia de produção mais freqüente de /l/ e /r/ intervocálicos é que determina em qual etapa do caminho proposto da aquisição o informante se encontra. Sabemos que o caminho da aquisição é uma tarefa complexa que envolve o amadurecimento de outras áreas, tais como a neuromotora, a cognitiva e a perceptual.

2.1. Dos informantes

Os informantes desta pesquisa são quatro (4) crianças, sendo três meninos e uma menina, na faixa etária dos 4:1 aos 7:4 anos de idade, que substituem /l/ e /r/ intervocálicos por [j]. Todos são monolíngües e pertencentes à classe social baixa.⁵ A triagem dos informantes foi realizada em escolas públicas e particulares e em clínicas de fonoaudiologia públicas e particulares. Tanto nas escolas quanto nas clínicas procuramos crianças com atraso na aquisição dos sons da fala. As crianças indicadas passaram por uma pequena triagem, onde mostramos oito figuras, quatro relacionadas ao /l/ intervocálico e as outras quatro relacionadas ao /r/ intervocálico, além de conversa espontânea. Aquelas crianças que, em algum momento, apresentaram a substituição de /l/ e/ou /r/ intervocálicos por [j], foram selecionadas como informantes da pesquisa.

Encontramos um pequeno número de informantes por se tratar de um fenômeno menos comum acima dos 4 anos de idade.⁶ O fato de termos encontrado mais meninos (3) que meninas (1) parece comum aos estudos de aquisição tardia. As variáveis sexo e classe social não serão, pois, levadas em consideração.

Trata-se de um estudo transversal (os dados são coletados por diferentes faixas de idade em um mesmo período), que, devido ao reduzido número de informantes conseguidos, tem caráter mais qualitativo que quantitativo, embora a quantidade de dados tenha permitido chegar a algumas informações estatísticas reveladoras de certos processos.

A seguir, o Quadro 1 traz as informações sobre cada informante, contendo: nome, sexo, idade, data de coleta dos dados e o total de palavras produzidas.

Nome	Sexo (anos:meses)	Idade	Data da coleta	Nº de palavras
Luísa	F	4:1	05/11/2001	80
José	M	5:1	16/10/2001	79
Baruc	M	6:10	19/11/2001	80
Abraão	M	7:4	09/11/2001	80

Quadro 1: Identificação dos Informantes

2.2. Da coleta de dados

Os dados foram coletados transversalmente, por meio de um instrumento proposto por Yavas, Hernandorena & Lamprecht (1992), acrescido de um desenho temático especificamente elaborado para eliciação das líquidas (AZAMBUJA, 1998), além de algumas figuras complementares do jogo Lince da Grow.

Em todos os instrumentos, levaram-se em consideração a posição da líquida na sílaba e na palavra, a tonicidade da sílaba em que se encontra a líquida, e a vogal co-ocorrente. Houve um quarto instrumento, utilizado para esta pesquisa que são três desenhos: uma “sala”, uma boneca chamada “Sara” e uma “saia”; utilizados para testar a consciência fonológica dos informantes na oposição entre o par /l/ e /r/ intervocálicos.

Para a coleta dos dados, todas as crianças foram gravadas na FALE/UFMG, em cabine isolada acusticamente, utilizando um gravador portátil de qualidade digital, da marca Sony, modelo TCD-D7, um microfone da marca Sony, modelo ECM-221 e fita DAT.⁷ Todos os dados foram obtidos através da nomeação espontânea e fala induzida. Foram mostradas as lâminas contendo desenhos, e o informante era induzido a dizer todas as palavras⁸ possíveis de eliciação das líquidas /l/ e /r/ intervocálicas, através da nomeação, ou de

perguntas ou de pequenas estórias contadas pelos informantes. No quarto instrumento proposto, além de ser solicitada a nomeação espontânea dos desenhos, era falado o nome de cada desenho, sem seguir uma ordem, e pedido ao informante para apontar o desenho solicitado. Este último instrumento serviu como apoio, complementando este estudo, visto que, não era nosso objetivo pesquisar, a fundo, a consciência fonológica dos informantes.

2.3. Da transcrição dos dados por informante

A transcrição fonética foi realizada utilizando-se o Alfabeto Fonético Internacional e a fonte SILDoulos IPA93. Só foram selecionadas para a transcrição as palavras contendo /l/ e /r/ intervocálicos.⁹ Através das transcrições, notamos que os informantes apresentavam outras substituições, além das substituições de /l/ e /r/ intervocálicos por [j] e de /r/ por [l]. Os processos que não envolveram /l/ e /r/ não foram analisados, pois não faziam parte do tema central desta pesquisa.

2.4. Do material coletado

O material de coleta dos dados desta pesquisa possui 76 palavras,¹⁰ sendo 47 ocorrências de /l/ e 33 ocorrências de /r/ para cada informante, o que perfaz, em tese, 320 dados.¹¹ A diferença entre o número de ocorrência das palavras com /l/ e /r/ intervocálicos se dá pelo fato de não existir, no Português, /r/ iniciando palavras, assim como não existe /r/ em monossílabos, restringindo-se, então, as suas possibilidades de ocorrências.

Para a análise dos dados, utilizamos o pacote de programas estatísticos Goldvarb 2001, muito empregado em análises lingüísticas variacionistas. Foram verificadas, para /l/ e /r/ intervocálicos, as produções mais freqüentes no *corpus*; as produções mais freqüentes por informante; e as influências da tonicidade, da vogal antecedente e seguinte, da posição na palavra e do número de sílabas em cada tipo de produção. A escolha das variáveis pesquisadas tem relação direta com as hipóteses formuladas durante a prática fonoaudiológica.

Mais detalhes sobre a análise propiciada pelo programa serão apresentados no decorrer da análise e descrição dos resultados, a seguir.

3. RESULTADOS

Nesta parte, serão apresentados os resultados obtidos, a partir da análise estatística propiciada pelo programa Goldvarb 2001. Os dados foram divididos em dois arquivos-chefe, um para o /l/ intervocálico e outro para o /r/ intervocálico, para serem processados separadamente pelo programa.

Antes da utilização do pacote Goldvarb, faz-se necessária, para cada arquivo-chefe, a definição das variáveis dependentes e independentes e dos fatores de cada variável. A variável dependente abarca as estratégias de produção dos segmentos pesquisados e está sempre presente para análise das variáveis independentes. As variáveis independentes são: os informantes, a tonicidade da sílaba em que o segmento se encontra, a vogal antecedente e seguinte ao segmento, a posição do segmento na palavra e o número de sílabas da palavra que contém o segmento. As variáveis independentes abarcam os fatores; por exemplo, a variável independente *informantes* contém os fatores: *José*, *Baruc*, *Abraão* e *Luísa*. Em seguida, um símbolo é atribuído a cada fator para a codificação dos dados, como por exemplo: **j** para o José, **b** para o Baruc, **a** para Abraão e **l** para a Luisa.

Estamos cientes de que temos poucos informantes e que o nosso material de coleta não propicia um número homogêneo de fatores das variáveis que serão analisadas; portanto, os percentuais e os índices de probabilidade obtidos de cada fator serão entendidos como indicadores de uma tendência ao favorecimento a cada estratégia de produções dos segmentos-alvo.

3.1. Variáveis e fatores do [l] intervocálico

A variável dependente, as variáveis independentes e seus fatores do [l] intervocálico foram determinados e codificados da seguinte forma:

Variável dependente:

- (0) – indica a realização do /l/: “vela” → [ˈvɛlə]
- (1) – indica a substituição de /l/ por [j]: “vela” → [ˈvɛjə]
- (2) – indica a não-realização de /l/: “vela” → [ˈvɛə]

Variáveis independentes:*a. Informantes:*

J – João **B** – Baruc **A** – Abraão **L** – Luisa

b. Tonicidade:

- e** – indica /l/ na sílaba pré-tônica: “geladeira”.
- t** – indica /l/ na sílaba tônica: “chocolate”.
- o** – indica /l/ na sílaba pós-tônica: “sala”.

c. Vogais antecedentes ao /l/ intervocálico:

- (**a**) – corresponde ao [a]: “balão” → [baˈlãu]
- (**e**) – corresponde ao [e]: “elefante” → [eleˈfãtɨ]
- (**i**) – corresponde ao [i]: “crocodilo” → [krokodɨˈlilɔ]
- (**o**) – corresponde ao [o]: “bolo” → [ˈbolu]
- (**u**) – corresponde ao [u]: “bule” → [ˈbuli]
- (**w**) – corresponde ao [ʊ]: “jaula” → [ˈʒaulə]
- (**c**) – corresponde ao [ɔ]: “bola” → [ˈbɔlə]
- (**g**) – corresponde ao [ɛ]: “panela” → [paˈnɛlə]
- (**∅**) – indica a ausência do segmento vocálico: “televisão” → [tɨˈsãu]

d. Vogais seguintes ao /l/ intervocálico:

- (**a**) – corresponde ao [a]: “gelado” → [ˈʒɛladu]
- (**e**) – corresponde ao [e]: “azulejo” → [azuˈlɛʒu]
- (**i**) – corresponde ao [i]: “o limão” → [u liˈmãu]
- (**o**) – corresponde ao [o]: “colorido” → [koloˈridu]
- (**u**) – corresponde ao [u]: “a luz” → [a ˈlus]
- (**w**) – corresponde ao [ʊ]: “cabelo” → [kaˈbɛlu]
- (**y**) – corresponde ao [ɪ]: “bule” → [ˈbulɨ]

- (c) – corresponde ao [ɔ]: “relógio” → [hɛ'lɔʒiu]
- (g) – corresponde ao [ɛ]: “picolé” → [piko'lɛ]
- (f) – corresponde ao [ə]: “janela” → [ʒa'hɛlə]
- (n) – corresponde ao [ã]: “melancia” → [mɛ'lãsiə]
- (ø) – indica a ausência do segmento vocálico: “televisão” → [tej'zãu]

e. *Posição na palavra:*

- (i) – indica que o /l/ intervocálico ocupa o início da palavra: “leite”
- (n) – indica que o /l/ intervocálico ocupa o meio da palavra: “pêlo”

f. *Número de sílabas:*

- (m) – indica que a palavra contém uma única sílaba: “luz”
- (d) – indica que a palavra contém duas sílabas: “laço”
- (t) – indica que a palavra contém três sílabas: “palito”
- (p) – indica que a palavra contém mais de três sílabas: “ensolarado”

3.1.1. Variável dependente: produções encontradas

Refletindo os dados obtidos, os percentuais e o número total de cada fator da variável dependente estão ilustrados na tabela abaixo.

/l/ > [l]	/l/ > [j]	/l/ > ø
79/187	61/187	47/187
42%	32%	25%

Tabela 1: Total em número e em percentual das produções do /l/ intervocálico

Desde que as realizações do segmento alvo representam a maior ocorrência no *corpus*, os percentuais, obviamente, serão mais altos neste caso. Interessa-nos, primordialmente, os dados referentes às outras produções e a correlação dos mesmos com o caminho proposto para a realização do /l/ intervocálico.

Analisaremos, agora, a partir dos dados, a interferência entre as variáveis dependente e independente.

3.1.2. A variável independente *informantes*

A análise da variável independente *informantes* (grupo A) permite especificar, por informante, qual produção do /l/ intervocálico é mais freqüente. Utilizamos os critérios de Azambuja (1998)¹² para indicar a fase da aquisição em que cada informante se encontra. Os critérios são os seguintes: a) menos de 50% de realização – o segmento ainda não está adquirido; b) de 51% a 85% de produções corretas – o segmento ainda está em fase de aquisição; c) de 86% a 100% – o segmento está efetivamente adquirido. Analisaremos, também, as produções mais freqüentes de acordo com a etapa em que cada informante se encontra no caminho proposto para a realização do /l/ intervocálico na aquisição com atraso.

3.1.2.1. Informante 1: José

O informante José apresenta as produções do /l/ intervocálico conforme a tabela a seguir:

/l/ > [l]	/l/ > [j]	/l/ > 0
0/46	18/46	28/46
0%	39%	61%

Tabela 2: Produções do /l/ intervocálico do José

O informante José se encontra na primeira etapa do caminho proposto para a realização do /l/ intervocálico. A tendência é que as não-realizações do /l/ intervocálico diminuam à medida que as substituições do /l/ intervocálico por [j] aumentem rumo à aquisição do segmento-alvo.

3.1.2.2. Informante 2: Baruc

A tabela 3 mostra as produções do /l/ intervocálico do informante Baruc:

/l/ > [l]	/l/ > [j]	/l/ > 0
42/47	4/47	0/47
92%	8%	0%

Tabela 3: Produções do /l/ intervocálico do Baruc

Conforme a tabela acima, o informante realiza o /l/ em 92% das suas produções e, portanto, já adquiriu efetivamente o /l/ intervocálico. O fato de haver uma pequena porcentagem de substituições por [j] confirma que o informante Baruc seguiu o caminho proposto para a realização do /l/ intervocálico.

3.1.2.3. Informante 3: Abraão

A tabela abaixo se refere às produções do /l/ intervocálico do Abraão:

/l/ > [l]	/l/ > [j]	/l/ > 0
10/47	26/47	11/47
21%	55%	24%

Tabela 4: Produções do /l/ intervocálico do Abraão

O informante Abraão ainda não adquiriu o /l/ intervocálico. Abraão se encontra na segunda etapa do caminho proposto para a realização do /l/ intervocálico. A tendência é de diminuir as não-realizações e as substituições por [j], à medida que as realizações do /l/ intervocálico aumentem rumo à sua aquisição efetiva.

3.1.2.4. Informante 4: Luísa

A tabela abaixo mostra as produções do /l/ intervocálico da informante Luísa:

/l/ > [l]	/l/ > [j]	/l/ > 0
29/47	6/47	12/47
62%	13%	25%

Tabela 5: Produções do /l/ intervocálico da Luísa

Luísa já se encontra na terceira etapa do caminho proposto para a realização do /l/ intervocálico, porém, ainda não adquiriu efetivamente o segmento-alvo.

Curiosamente, a informante Luísa contrasta com demais informantes por não seguir, de forma sistemática, o caminho proposto para a realização do /l/ intervocálico. Os dados parecem indicar que, em alguns casos, ela utiliza o caminho proposto para a realização do /l/ intervocálico na aquisição com atraso (passagem por [j]), mas que, em outros casos, ela não o utilizaria, passando direto da não-realização para a realização, como na aquisição normal.

3.1.3. As demais variáveis independentes do [l] intervocálico

As outras variáveis independentes (*tonicidade, vogais antecedentes, vogais seguintes, posição na palavra, número de sílabas*) apresentaram, respectivamente, os seguintes resultados:¹³

- Há uma maior dificuldade de se atingir o alvo na posição pós-tônica do que nas outras posições. Este fato, substituição de /l/ por [j] na sílaba pós-tônica, também foi observado por Azambuja (1998).
- Foi constatada uma tendência ao favorecimento da substituição por [j] pelas vogais médias e anteriores [ɛ] e [e], o que equivale a dizer que esses ambientes vocálicos antecedentes dificultam a realização do som alvo. Azambuja (1998) constatou que [ɛ], em contexto antecedente, promove a substituição por [j]. Observamos que é comum a inserção de [j] entre duas vogais, principalmente se a primeira vogal é [ɛ] ou [e], como no nome “Léa” que é produzido como “Léia” ou na palavra “mear”, dita “meiar”.

- A não-realização do /l/ intervocálico por [j] é mais freqüente tendo como contexto antecedente as vogais altas e posteriores [u] e [u]. Neste caso, os dados apontam para uma tendência a uma passagem direta da não-realização para o som alvo, como na aquisição normal.
- A substituição do /l/ intervocálico por [j] tende a se favorecer diante do contexto seguinte representado pelas vogais [u] com 66% das ocorrências [ə] com 59%, [i]¹⁴ com 53% e [o] com 42%. O fato de as vogais pós-tônicas finais [u] e [ə] serem as mais favoráveis à substituição por [j] era esperado, uma vez que estas vogais são as que normalmente aparecem em sílaba pós-tônica. Estas vogais, nesse contexto, parecem segurar os informantes por mais tempo na segunda etapa rumo à realização do /l/ intervocálico. A vogal média e anterior [e] tende a favorecer a não-realização do [l] intervocálico, dificultando ainda mais a passagem rumo ao segmento-alvo.
- A variável independente *posição na palavra* refere-se às posições iniciais e não-iniciais do /l/ intervocálico na palavra. A posição inicial retrata a posição de *onset* inicial¹⁵ e a posição não-inicial retrata a posição de *onset* medial. O fato de o ambiente do /l/ ser intervocálico causa a impressão de que todo /l/ está em posição não-inicial ou *onset* medial; por isso, resolvemos observar se os informantes discriminam estas duas posições e se comportam diferentemente em suas produções. Constatamos que a substituição do /l/ intervocálico por [j] tende a acontecer mais em posição não-inicial na palavra, fenômeno esse também observado por Azambuja (1998). No entanto, no percurso para a realização do /l/ intervocálico por [j] em posição inicial na palavra, o informante evita a substituição por [j]. Os dados parecem demonstrar que há uma consciência, por parte dos informantes, de que se trata de uma construção morfológica diferente do /l/ puramente intervocálico, como, por exemplo, em “Alan” x “a lâ”.
- A variável independente *número de sílabas* não se mostrou significativa ao apontar o favorecimento de cada produção do /l/ intervocálico em relação ao número de sílabas, visto que, há

o predomínio das realizações do /l/ intervocálico em todos os tamanhos de palavra. A diferenciação das palavras monossílabas e dissílabas foi importante somente para apontar que as monossílabas favorecem muito pouco a substituição do /l/ intervocálico por [j]. Aqui também, como na variável anterior, os dados parecem demonstrar que há uma consciência, por parte dos informantes, de que se trata de uma construção morfológica diferente do /l/ puramente intervocálico; portanto, a passagem por [j] parece não ser um caminho obrigatório para aquisição do som alvo, aproximando-se da aquisição normal.

3.1.4. Variáveis estatisticamente relevantes para a realização do [l] intervocálico

Nenhuma variável independente ou grupo de fatores foi selecionado como estatisticamente significativo. Isto quer dizer que nenhuma combinação das variáveis ou dos grupos de fatores foi relevante para apontar o favorecimento à realização do /l/ intervocálico.

3.1.5. Variáveis estatisticamente relevantes para a substituição do [l] intervocálico por [j]

O programa elegeu a combinação das variáveis *vogais seguintes* e *posição na palavra* como estatisticamente relevantes para substituição do /l/ intervocálico por [j]. Esta combinação obteve significância de **0.001**, o que equivale dizer que a margem de erro na análise probabilística do programa é igual a zero. O programa ainda mostra uma hierarquia de relevância entre as variáveis selecionadas, sendo o grupo de fatores das *vogais seguintes* o “grupo mais forte”. A vogal seguinte [i] é a que mais influencia na substituição do /l/ intervocálico por [j]; já a vogal seguinte [a] é pouco favorável a esta produção. Assim, com os índices de probabilidade, o contexto vocálico seguinte, mais favorável é a vogal [i] com índice de probabilidade de 0.839, seguida das vogais [ə], [u] e [o] com índices de probabilidade

em torno de 0.6. Já os percentuais foram mais altos para as vogais finais [u] com 66% das ocorrências, [ə] com 59%, [i] com 53% e [o] com 42%.

Consideramos que os resultados provenientes dos índices de probabilidades, e não o dos percentuais, são mais condizentes ao favorecimento da substituição do /l/ intervocálico por [j], visto que [i] é a vogal que mais compartilha semelhanças com [j].

A razão para que a vogal [a] seja, de todas as vogais, a que menos favorece a substituição do /l/ intervocálico por [j] pode relacionar-se ao fato de ser a vogal mais baixa, distante, pois, da produção do [j].

Na variável *posição na palavra*, a posição inicial do segmento favorece muito pouco à substituição do /l/ intervocálico por [j], apresentando o índice de probabilidade de **0.254**. Parece-nos que, no percurso para a realização do intervocálico /l/ em posição inicial na palavra, o informante tende a passar diretamente da não-realização para a realização do segmento-alvo, devido à consciência sobre construção morfológica diferente.

3.1.6. Variáveis estatisticamente relevantes para a não-realização do /l/ intervocálico

O programa elegeu a combinação das variáveis *vogais antecedentes* e *vogais seguintes* como estatisticamente relevantes para a não-realização do /l/ intervocálico. Esta combinação obteve significância de **0.019**. O programa selecionou o grupo de fatores das *vogais seguintes* como o “grupo mais forte”. Das *vogais antecedentes*, as altas posteriores são as mais favoráveis à não-realização do /l/ intervocálico e a vogal menos favorável é a vogal média posterior [o]. Entre as vogais altas e posteriores, [u] é a mais atingida pela não-realização do /l/ intervocálico, apresentando índice de probabilidade de 0.857. A vogal [o] é a menos favorável à não-realização do /l/ intervocálico, apresentando o menor índice de probabilidade de 0.201.

Consideramos que as vogais antecedentes altas posteriores tendem a inviabilizar a realização do /l/ intervocálico por este ser um segmento alto e anterior. Os dados apontam para uma tendência a uma passagem direta da não-realização para o som alvo.

Das *vogais seguintes*, notamos que a vogal mais favorável à não-realização do /l/ intervocálico é a vogal média e posterior [ɔ] apresentando o índice de probabilidade de 0.855 e, em seguida, a ausência do segmento vocálico [ø] com índice de 0.843.

A variável vogais seguintes *mostrou* ser relevante para as produções desviantes¹⁶ do /l/ intervocálico: substituição por [j] e não-realização. Comparando os resultados obtidos em probabilidades, as vogais menos favoráveis à não-realização do /l/ intervocálico, que são [i] com índice de 0.187 e [u] com índice de 0.195, são as mais favoráveis à substituição do /l/ intervocálico por [j].

3.2. Descrição e análise dos dados do /r/ intervocálico

As produções do /r/ intervocálico encontradas foram: substituição do /r/ intervocálico por [l]; substituição do /r/ intervocálico por [j] e não-realização do /r/ intervocálico. Nenhum dos quatro informantes, em algum momento, realizou o /r/ intervocálico. Então, a variável dependente, as variáveis independentes e os fatores do /r/ intervocálico foram determinados e codificados da seguinte forma:

Variável dependente:

- (0) – indica a substituição de /r/ por [l]: “Sara” → [ˈsalə]
- (1) – indica a substituição de /r/ por [j]: “Sara” → [ˈsajə]
- (2) – indica a não-realização de /r/: “Sara” → [ˈsaə]

Variáveis independentes:

a. *Informantes*:

J – João **L** – Luísa **A** – Abraão **B** – Baruc

b. *Tonicidade*¹⁷:

- t** – indica /r/ na sílaba tônica: “morango”.
- o** – indica /r/ na sílaba pós-tônica: “geladeira”.

c. *Vogais antecedentes ao /r/:*

- (a) – corresponde ao [a]: “barata” → [ba¹ratə]
- (e) – corresponde ao [e]: “pêra” → [p¹erə]
- (i) – corresponde ao [i]: “direito” → [dʒi¹rejtʊ]
- (o) – corresponde ao [o]: “vassoura” → [va¹sorə]
- (u) – corresponde ao [u]: “buraco” → [bu¹rakʊ]
- (s) – indica a ausência do segmento vocálico [ø]: “árvore” → [l¹avi]

d. *Vogais seguintes ao /r/¹⁸:*

- (a) – corresponde ao [a]: “parada” → [pa¹radə]
- (e) – corresponde ao [e]: “direito” → [dʒi¹rejtʊ]
- (i) – corresponde ao [i]: “nariz” → [na¹ris]
- (o) – corresponde ao [o]: “dinheiro” → [dʒi¹nejrʊ]
- (w) – corresponde ao [u]: “dinheiro” → [dʒi¹nejrʊ]
- (y) – corresponde ao [I]: “árvore” → [l¹ahvorɪ]
- (g) – corresponde ao [ɛ]: “jacaré” → [ʒaka¹rɛ]
- (f) – corresponde ao [ə]: “cadeira” → [ka¹dejrə]
- (n) – corresponde ao [ã]: “aranha” → [a¹rãɲə]

e. *Número de sílabas¹⁹:*

- (d) – indica que a palavra contém duas sílabas: “pêra”
- (t) – indica que a palavra contém três sílabas: “cenoura”
- (p) – indica que a palavra contém mais de três sílabas: “saboneteira”

Diferentemente do /l/, a variável *posição na palavra* não foi pesquisada porque não existe /r/ iniciando vocábulos. Todas as palavras contêm /r/ não-inicial.

Passaremos, agora, a analisar os resultados das variáveis dependentes.

3.2.1. Variável dependente

A variável dependente reflete as produções encontradas na aquisição com atraso do /r/ intervocálico, respectivamente: **0**, **1** e **2**, correspondentes a /r/ > [l], /r/ > [j] e /r/ > 0, aqui denominadas de

substituição do /r/ intervocálico por [l], substituição do /r/ intervocálico por [j] e não-realização do /r/ intervocálico.

A tabela abaixo mostra os percentuais e o número total de cada fator da variável dependente:

/r/ > [l]	/r/ > [j]	/r/ > 0
21/132	86/132	17/132
16%	69%	13%

Tabela 6: Total em número e percentual das produções do /r/

Observando a tabela 6, notamos que o arquivo do /r/ intervocálico contém 132 dados obtidos das produções dos quatro informantes. De acordo com o número total dos dados, 21 palavras (16%) apresentaram a substituição /r/ intervocálico por [l], 86 palavras (69%) tiveram /r/ intervocálico substituído por [j] e somente 17 palavras (13%) apresentaram a não-realização do /r/ intervocálico. No arquivo do /r/ intervocálico, a substituição do /r/ intervocálico por [j] é, de longe, a produção mais freqüente.

Desde que as substituições do /r/ intervocálico por [j] representam a maior ocorrência no *corpus*, esperamos encontrar, em cada fator das variáveis independentes, o maior número de ocorrências de substituição por [j], seguida da substituição por [l] e, por último, com o menor número de ocorrências possíveis, a não-realização. Desta forma, o caminho proposto para a realização do /r/ intervocálico é seguido, mesmo diante de uma permanência temporária na segunda etapa da aquisição (conferir item 2). Interessa-nos, primordialmente, analisar os fatores que apontem diferenças no caminho proposto para a realização do /r/ intervocálico na aquisição com atraso.

A seguir, analisaremos, em percentuais, a interferência entre as variáveis dependente e independentes.

3.2.2. A variável independente informantes

A análise da variável *sujeitos* (grupo A) permite especificar, por informante, qual é a produção do /r/ intervocálico mais freqüente.²⁰

Analisaremos as produções mais frequentes, definindo a etapa²¹ em que cada informante se encontra no caminho proposto para a realização do /r/ intervocálico.

3.2.2.1. Informante 1: José

O informante José apresenta as produções do /R/ intervocálico conforme a tabela a seguir:

/r/ > [ɹ]	/r/ > [j]	/r/ > ɒ
0/33	25/33	8/33
0%	76%	24%

Tabela 7: Produções do /r/ intervocálico do José

A tabela 7 nos mostra que as substituições do /r/ intervocálico por [j] são muito frequentes, ocorrendo em 61% das produções do /r/ intervocálico do José. A tabela 16 revela também que substituição do /r/ intervocálico por [ɹ] não acontece, visto que o informante ainda não realiza o /l/ intervocálico, conforme vimos na tabela 2, item 3.1.2.1. O informante José se encontra na segunda etapa do caminho rumo à realização do /r/ intervocálico.

3.2.2.2. Informante 2: Baruc

A tabela 8 mostra as produções do /r/ intervocálico do informante Baruc:

/r/ > [ɹ]	/r/ > [j]	/r/ > ɒ
5/33	25/33	3/33
15%	76%	10%

Tabela 8: Produções do /r/ intervocálico do Baruc

Embora o informante já tenha adquirido efetivamente o /l/ intervocálico (confira a tabela 3 no item 3.1.2.2.) ele se detém na segunda etapa do caminho rumo à realização do /r/ intervocálico,

onde há o predomínio das substituições por [j]. Esperávamos encontrar o predomínio das substituições por [l] nas produções do /r/ intervocálico, porém, o que encontramos foram 76% das ocorrências de substituições por [j] e 15% por [l]. Há duas explicações para tal fato: a primeira, é que o informante segue o caminho proposto para a realização com atraso do /r/ intervocálico e tenderá a diminuir as substituições do /r/ intervocálico por [j], à medida que as substituições do /r/ intervocálico por [l] aumentarem rumo à realização do segmento alvo; a segunda, é que o Baruc, tendo consciência de que /r/ e /l/ são fonemas distintos, evita a substituição de /r/ intervocálico por [l] e, por isso, não segue sistematicamente o caminho proposto, devendo passar diretamente da substituição por [j] para a realização do /r/ intervocálico.

3.2.2.3. Informante 3: Abraão

A tabela abaixo refere-se às produções do /r/ intervocálico do Abraão:

/r/ > [l]	/r/ > [j]	/r/ > 0
0/33	32/33	1/33
0%	97%	3%

Tabela 9: Produções do [r] intervocálico do Abraão

De acordo com a tabela 9, o informante Abraão, praticamente, só substitui /r/ intervocálico por [j] abarcando 97% das ocorrências de suas produções. A substituição do /r/ intervocálico por [l] ainda não acontece porque o Abraão ainda não adquiriu o /l/ intervocálico (confira a tabela 4 no item 3.1.2.3.). Sendo assim, o informante se encontra na segunda etapa do caminho proposto para o /r/ intervocálico e a tendência é que as substituições do /r/ intervocálico por [j] diminuam, à medida que as substituições do /r/ intervocálico por [l] aumentem rumo à realização do segmento-alvo, ou que ele passe diretamente de [j] para [r].

3.2.2.4. Informante 4: Luísa

A tabela abaixo mostra as produções do /r/ intervocálico da informante Luísa:

/r/ > [ɹ]	/r/ > [j]	/r/ > ɒ
17/33	11/33	5/33
52%	33%	15%

Tabela 10: Produções do /r/ intervocálico da Luísa

Observando a tabela 10, vemos que a substituição do /r/ intervocálico por [ɹ], com 52% das ocorrências, é a produção mais freqüente da informante Luísa. Tal fato é explicado porque a informante se encontra em fase de aquisição do /l/ intervocálico (confira a tabela 5 no item 3.1.2.4.). Luísa se encontra na terceira etapa ou penúltima etapa no caminho proposto para o /r/ intervocálico. Pelo visto, esta informante não passará diretamente da produção [j] para [r], seguindo à risca cada etapa.

3.2.3. A variável independente *tonicidade*

A tabela 11 mostra a relação entre a tonicidade e as produções do /r/:

Grupo B	/r/ > [ɹ]	/r/ > [j]	/r/ > ɒ	Total
T	8/60 13%	42/60 70%	10/60 16%	48%
O	13/64 20%	44/64 68%	7/64 10%	51%

Tabela 11: Tonicidade versus produções do [r] intervocálico

Observando a variável independente *tonicidade*, notamos que, em posição tônica, as não-realizações são um pouco mais freqüentes que as substituições por [ɹ], o que indica que, nesta posição, os informantes podem evitar a terceira etapa do caminho proposto para a realização do /r/ intervocálico, passando direto da substituição por [j] à realização do alvo. Os estudos sobre aquisição do /r/, quando

comparados, apresentam achados contraditórios sobre o favorecimento da tonicidade à realização do [r]. Miranda (1996) observou que a posição tônica é mais favorável à realização do /r/, e Hernandorena e Lamprecht (1997) observaram ser a posição pós-tônica a mais favorável à realização do /r/. Contudo, o estudo de Hernandorena e Lamprecht (1997) possui mais informantes além de ter sido feita a análise estatística dos dados.

3.2.4. A variável independente vogais antecedentes

A tabela 12 apresenta a influência das vogais antecedentes nas produções do /r/:

Grupo C	/r/ > [l]	/r/ > [j]	/r/ > 0	Total
[a]	8/58 13%	45/58 77%	5/58 8% 0	46%
[e]	6/25 24%	18/25 72%	1/25 4%	20%
[i]	0	0	4/4 100%	4%
[o]	6/24 25%	14/24 58%	4/24 16%	19%
[u]	1/8 12%	5/8 62%	2/8 25%	6%
[ø]	0	0	5/5 100%	4%

Tabela 12: Vogais antecedentes versus produções do /r/ intervocálico

Podemos constatar, na tabela acima, que a vogal antecedente [i] e a ausência do segmento vocálico antecedente [ø] são favoráveis à não-realização do /r/ intervocálico e tendem a dificultar a passagem às outras etapas rumo ao segmento-alvo. A vogal [u] apresenta percentuais mais altos de não-realização que os de substituição por [l], indicando que o informante pode evitar a terceira etapa, diante deste ambiente vocálico antecedente, passando diretamente da substituição por [j] à realização do /r/ intervocálico. De modo geral, podemos observar uma estagnação na produção [j], o que pode ser indício de se evitar o [l], por considerá-lo fonologicamente diferente de /r/.

3.2.5. A variável independente vogais seguintes

A tabela 13 mostra a ocorrência de cada produção do /r/ intervocálico de acordo com o contexto vocálico seguinte.

Grupo D	/r/ > [l]	/r/ > [j]	/r/ > 0	Total
[a]	4/23 17%	14/23 45%	5/23 21%	8%
[e]	¼ 25%	¾ 75%	0	3%
[i]	0	16/16 100%	0	12%
[o]	0	2/2 100%	0	1%
[u]	1/6 16%	4/6 66%	1/6 16%	4%
[ɪ]	0	0	4/4 100%	3%
[ɛ]	0	4/6 66%	2/6 33%	4%
[ə]	12/52 23%	38/52 73%	2/52 3%	41%
[ã]	3/11 27%	5/11 45%	3/11 27%	8%

Tabela 13: Vogais seguintes versus produções do /r/ intervocálico

Analisando a variável independente vogais seguintes, notamos que, diante da vogal [a], o número de não-realizações é maior que o número de substituições por [l] indicando que, nesse contexto, os informantes podem evitar a terceira etapa do caminho proposto para alcançar o /r/ intervocálico. A vogal [i], que é a segunda vogal mais freqüente do *corpus*, mostrou-se favorável à substituição por [j], com 100% das ocorrências, uma vez que consideramos todas as produções de [i] como [ji], onde [j] representa a substituição de /r/, visto que, são mais freqüentes as substituições por [j] diante de outros ambientes vocálicos seguintes. A vogal [ɪ] mostrou-se favorável à não-realização, com 100% das ocorrências, porque, na palavra “árvore”, todos os informantes não realizaram o /r/ intervocálico e também outros segmentos da palavra. Diante das vogais [i], [o] e [ɛ] há o predomínio das substituições por [j], o que representa uma estagnação na segunda etapa do caminho proposto para a realização do /r/ intervocálico na aquisição com atraso, indício de consciência fonológica da distinção entre /l/ e /r/, facilitando a passagem direta de [j] para /r/.

3.2.6. A variável independente número de sílabas

A tabela 14 mostra a relação entre o número de sílabas e as produções do /r/ intervocálico:

Grupo D	/r/ > [l]	/r/ > [j]	/r/ > 0	Total
Dissílaba	2/12 16%	10/12 83%	0	9%
Trissílaba	14/84 16%	53/84 63%	17/84 20%	67%
Polissílaba	5/28 17%	23/28 82%	0	22%

Tabela 14: Número de sílabas versus produções do /r/ intervocálico

Conforme a tabela acima, notamos que a maior parte das palavras do *corpus* é trissílaba, com 67% da ocorrência total, e este é o ambiente que favorece menos a substituição do /r/ intervocálico por [j], com 63 % contra 83% das ocorrências em palavras dissílabas e 82% em palavras polissílabas. Observamos, também, que nas palavras trissílabas o número de não-realizações é um pouco maior que o número de substituições por [l], o que indica que os informantes podem optar, nesse contexto, por evitar a terceira etapa do caminho proposto e passar diretamente da substituição por [j] ao segmento-alvo. De qualquer forma, o reforço na realização [j] em todos os casos pode apontar para uma passagem direta para [r], nesses casos.

Fizemos o cálculo das probabilidades para verificar, com mais clareza, a interferência entre as variáveis. Subdividimos o arquivo do /r/ intervocálico em três arquivos: o primeiro arquivo apresentou a análise das variáveis estatisticamente relevantes para a substituição do /r/ intervocálico por [l]; o segundo, a análise das variáveis estatisticamente relevantes para a substituição do /r/ intervocálico por [j]; e o terceiro, a análise das variáveis estatisticamente relevantes para a não-realização do /r/ intervocálico. Como já se esperava, o programa não apontou nenhuma variável estatisticamente relevante para qualquer produção do /r/ intervocálico. Foram tantos os *knockouts* que algumas variáveis inteiras foram retiradas e, para outras, restaram poucos fatores para análise. A razão deste fato é a mesma encontrada nos percentuais: no *corpus* encontramos a maior ocorrência de substituição do /r/ intervocálico por [j], onde todos os

fatores apresentam-se favoráveis a esta produção, que parece uma produção manifestadora da consciência fonológica dos informantes ao realizarem /l/ para [l] e manterem [j] para /r/.

O que se observa no caminho proposto para a realização do /r/ intervocálico é que há uma estagnação muito grande na segunda etapa. Isto, repetimos, parece uma estratégia adotada pelos informantes para, em princípio, não usarem o [l], que tem valor fonêmico diferente de /r/.

Concluimos que a terceira etapa da aquisição com atraso do /r/ intervocálico é, por vezes, opcional, mas defendemos que a aquisição definitiva do [r] depende da aquisição do /l/ visto que: ambos foram substituídos por [j]; a maioria dos fatores apresentou, mesmo que em baixa ocorrência, a substituição de /r/ por [l]; não encontramos o /r/ adquirido antes do /l/.

4. CONCLUSÕES ACERCA DAS HIPÓTESES LEVANTADAS

A partir da comparação dos resultados obtidos neste estudo com as hipóteses previamente formuladas, as quais serão apresentadas a seguir, é possível chegar-se a um resumo das conclusões acerca da realização de /l/ e /r/ intervocálicos na aquisição com atraso.

a) O /l/ é realizado antes de /r/.

Observando as tabelas 7 a 10, do item 3.2.2., notamos que todos os informantes não realizaram o /r/ sequer uma vez, mesmo aqueles que já tivessem realizado o /l/ (confira tabelas do item 3.1.2.), o que confirma a nossa hipótese.

b) /l/ e /r/ podem ser não-realizados, substituídos ou realizados.

Como vimos na hipótese anterior, somente a realização do /l/ acontece; fora isso, as outras estratégias de produção manifestam-se em ambas as líquidas. Observamos que a realização é a produção mais freqüente para o /l/ (confira a tabela 1 do item 3.1.1.) e a substituição por [j] é a mais freqüente para o /r/ (confira a tabela 6 do item 3.2.1.). A hipótese foi parcialmente confirmada, pois não ocorreu a realização de /r/, embora possamos prever que ela irá acontecer em determinado momento.

c) Quando encontramos /l/ substituído por [j], encontramos também o /r/ substituído por [j].

Todos os informantes substituíram /l/ e /r/ por [j] em algum contexto (confira as tabelas dos itens 3.1.2 e 3.2.2.).

d) *Se o informante já produz o /l/, ele não irá mais substituir o /r/ por [j], e sim por [l].*

Os resultados obtidos não confirmam esta hipótese e apontam a substituição de /r/ por [j] como a produção mais freqüente da maioria dos informantes, com exceção da Luísa, que apresenta a substituição de /r/ por /l/ como a mais freqüente (confira a tabela 5, do item 3.1.2., e 10 do item 3.2.2.4.). A estagnação na substituição de /R/ por [j] e a pouca ocorrência da substituição de /r/ por [l], reafirma a consciência fonológica da criança, que distingue /l/ de /r/, e, por isso, mantém a oposição evitando a substituição por [l] para o fonema /r/.

e) A posição pós-tônica é a mais atingida pela substituição de /l/ e /r/ por [j].

Esta hipótese se confirma apenas para a líquida não-lateral, já que a maior parte das substituições de /l/ por [j] (58%) concentram-se na posição pós-tônica (confira 3.1.3.). Quanto à líquida não-lateral, os resultados apontam a posição tônica como a mais atingida pelas substituições de /r/ por [j] (70%), embora a posição pós-tônica tenha sido, também, muito atingida (68%) (confira a tabela 11, no item 3.2.3.). Este resultado deve estar relacionado ao que comentamos na hipótese anterior. O pacote de programas estatísticos Goldvarb 2001 não selecionou a variável *tonicidade* como estatisticamente relevante para nenhuma estratégia de produção de /l/ e /r/ intervocálicos.

f) *Tanto as vogais antecedentes quanto seguintes exercem influências nas estratégias de produção do /l/ e /r/.*

O Goldvarb 2001 apontou a variável *vogais antecedentes* como estatisticamente relevante somente para a não-realização do /l/, e a variável *vogais seguintes* para a não-realização do /l/ e para a substituição do /l/ por [j]. A análise da influência das vogais antecedentes possibilitou destacar as vogais posteriores [ʊ] e [u] como as mais

favoráveis à não-realização do /l/. Os resultados obtidos das vogais seguintes apontam a vogal [ɔ] e a ausência do segmento vocálico [ø] como os ambientes seguintes mais favoráveis à não-realização do /l/, e a vogal [i] para a substituição do /l/ por [j].

g) Encontramos mais substituições por [j] em posição não-inicial na palavra.

Os resultados obtidos confirmam esta hipótese e apontam a posição não-inicial do /l/ como a mais favorável à substituição por [j]. Entretanto, confira a nota 19 para o /r/ inicial.

b) *Quanto maior a palavra maior o número de não-realizações de /l/ e /r/.*

Esta hipótese não se confirma, pois, mesmo que a não-realização do /l/ tenha atingido o seu maior número de ocorrências em palavras polissílabas, o número de realizações é, ainda, superior. Não houve a não-realização do /r/ em palavras polissílabas (confira a tabela 14, no item 3.2.6.).

NOTAS

¹ Utilizamos a palavra “produção” com o sentido de “como produziu”.

² A não-realização é uma forma de produção, considerada como produto.

³ Na verdade, tanto as não-realizações, as substituições e a própria realização de /l/ e /r/ intervocálicos estão no nível fonético, pois a criança já possui a distinção /l/ e /r/ no nível fonológico. Portanto, não se trata realmente de “aquisição fonológica”, mas da descoberta de um caminho para a produção dos fonemas /l/ e /r/. Por isso, ao longo deste trabalho, o termo “aquisição” estará, freqüentemente, sendo usado como sinônimo de “realização”.

⁴ Esse objetivo não será alvo de discussão nesse artigo. Para maiores detalhes, cf. MARTINS (2002).

⁵ De acordo com o IBGE, a classe média-baixa é aquela cuja renda familiar gira em torno de 2 a 5 salários mínimos. Embora tenhamos percorrido escolas e clínicas que abarcassem todos os níveis sociais, todos os informantes encontrados são de classe social baixa.

⁶ Segundo Lamprecht (1997), a maioria dos processos envolvidos na aquisição de segmentos isolados desaparecem até os 4;1ou 4;2.

⁷ Os cuidados com o ambiente acústico, desprovido de ruído, foram tomados, visando à qualidade sonora para a melhor transcrição dos dados e para a análise acústica pretendida *a priori*.

⁸ Refere-se a substantivos, verbos e adjetivos.

⁹ Essas palavras podem conter o segmento [l] em posição não-inicial, como em “balão”, ou em posição inicial, como em “lobo”. Neste último exemplo, todos os informantes foram induzidos a usarem o mesmo artigo (ex.: “o lobo”), criando, desta forma, o ambiente intervocálico. O segmento [r] aparece apenas em posição não-inicial na palavra (ex.: “pêra”).

¹⁰ A lista é composta de 76 palavras, porém 4 palavras possuem tanto /l/ quanto /r/ intervocálicos e foram contadas duas vezes; daí, o total de 80 ocorrências.

¹¹ Na verdade, o total de dados foi de 319, porque o informante José não produziu uma das palavras.

¹² Azambuja fez uma adaptação de Hernandorena (1990) que propõe 4 estágios para a aquisição dos segmentos. Azambuja propôs 3 estágios por considerar a diferença entre eles muito sutil para influenciar os resultados.

¹³ Não serão detalhados aqui os gráficos e tabelas obtidos nessa análise. Para uma visão detalhada dos mesmos, cf. MARTINS (2002), cap.3.

¹⁴ Todas as produções de [il] foram consideradas como [jl], onde [j] representa a substituição de /l/.

¹⁵ Independente da presença de um artigo.

¹⁶ “produções desviantes” não têm o mesmo sentido de “aquisição com desvios”.

¹⁷ O material de coleta não continha palavras com /r/ pré-tônico como, por exemplo, “caridade”.

¹⁸ As produções fonéticas, contendo /r/, tanto nas vogais antecedentes quanto seguintes, são meros exemplos que não constam dos dados obtidos.

¹⁹ Não existe /r/ em posição de *onset* inicial somente em posição de *coda*, que não faz parte desta pesquisa. Também não existe em Português vogal nasal antecedendo /r/, pois, neste caso, será sempre [r] (confira CÂMARA JÚNIOR, 1970).

²⁰ Não foram utilizados os critérios de Azambuja (1998) citados no item 3.1.2., pois nenhum informante realizou o /r/ intervocálico.

²¹ O caminho proposto para a aquisição do /r/ ocorre em 4 etapas (conferir item 2).

REFERÊNCIAS BIBLIOGRÁFICAS

- AZAMBUJA, E. J. M. *A aquisição das líquidas laterais do português: um estudo transversal*. 1998. Dissertação (Mestrado em Letras) - Instituto de Letras e Artes da PUCRS, Porto Alegre.
- CÂMARA Jr., J. M. *Estrutura da língua portuguesa*. Rio de Janeiro: Vozes, 1970.
- CLEMENTS, G. N. *The geometry of phonological features*. *Phonology Yearbook*, v.2, p. 225-252, 1985.
- CLEMENTS, G. N.; HUME, E. V. The internal organization of speech sounds. In: GOLDSMITH, J. E. (Org.) *The handbook of phonological theory*. London: Basil Blackweel, 1995.
- HERNANDORENA, C. L. M. *A aquisição da fonologia do português: estabelecimento de padrões com base nos traços distintivos*. 1990. Tese (Doutorado em Lingüística) - Instituto de Letras e Artes da PUCRS, Porto Alegre.
- HERNANDORENA, C. L. M. Relações implicacionais na aquisição da Fonologia. *Letras de Hoje*. Porto Alegre. v.31, n. 2, p.67-76, jun. 1996.
- HERNANDORENA, C. L. M. Aquisição da fonologia e implicações teóricas: um estudo sobre as soantes palatais. In: LAMPRECHT, R. R. (Org.). *Aquisição da Linguagem: questões e análises*. Porto Alegre, 1999.
- HERNANDORENA, C. L. M. Restrições segmentais e prosódicas na aquisição das líquidas do português brasileiro e do português europeu. In: *II Congresso Internacional da ABRALIN*, 2001, Fortaleza.
- HERNANDORENA, C. L. M.; LAMPRECHT, R. R. A aquisição das consoantes líquidas do Português. *Letras de Hoje*, Porto Alegre. v. 32, n. 4, p.7-22, dez. 1997.
- LAMPRECHT, R. R. *Perfil da aquisição normal da fonologia normal do português – descrição longitudinal de 12 crianças: 2:9 a 5:5*. 1990. Tese (Doutorado em Lingüística) - Instituto de Letras e Artes da PUCRS, Porto Alegre.
- LAMPRECHT, R. R. Aquisição da fonologia do Português na faixa etária dos 2:9 aos 5:5. *Letras de Hoje*. Porto Alegre: EDIPUCRS, v. 28, n. 2, p.99-106, 1997.
- LAMPRECHT, R. R. (Org.) Desvios fonológicos: evolução nas pesquisas, conhecimento atual e implicações dos estudos em Fonologia Clínica. In: LAMPRECHT, R. R. (Org.). *Aquisição da Linguagem: questões e análises*. Porto Alegre: EDIPUCRS, 1999.
- LAMPRECHT, R. R. As líquidas não-laterais na aquisição do Português Brasileiro – estudo comparativo entre o desenvolvimento fonológico normal e os desvios fonológicos evolutivos. In: *II Congresso Internacional da ABRALIN*, 2001, Fortaleza.

- LOWE, R. J. *Fonologia – Avaliação e intervenção: Aplicações na patologia de fala*. Trad. Marcos A. G. Domingues. Porto Alegre: Artes Médicas, 1996. (original inglês)
- MARTINS, M. *O caminho para a realização de /l/ e /r/ intervocálicos na aquisição com atraso: a substituição por [jl]*. 2002. Dissertação (Mestrado em Linguística) - Faculdade de Letras da UFMG, Belo Horizonte.
- MIRANDA, A. R. M. *A aquisição do 'r': uma contribuição à discussão sobre seu status fonológico*. 1996. Dissertação (Mestrado em Letras) - Instituto de Letras e Artes da PUCRS, Porto Alegre.
- MIRANDA, A. R. M. A aquisição das líquidas não-laterais no português do Brasil. *Letras de Hoje*. Porto Alegre. V.33, n. 2, p. 123-131, junho 1998.
- MIRANDA, A. R. M. /Kar.ro/ ou /ka.Ro/: evidências da aquisição da linguagem. In: II Congresso Internacional da ABRALIN, 2001, Fortaleza.
- MOTA, H. B. *Aquisição segmental do Português: um modelo implicacional de complexidade de traços*. 1996. Tese (Doutorado em Letras) - Instituto de Letras e Artes da PUCRS, Porto Alegre.
- MOTA, H. B. Aquisição segmental do Português: um modelo implicacional de complexidade de traços. *Letras de Hoje*. Porto Alegre. v. 32, n. 4, p. 23-47, dez. 1997.
- MOTA, H. B. Os caminhos na aquisição segmental do Português. In: LAMPRECHT, R. R. (Org.). *Aquisição da Linguagem: questões e análises*. Porto Alegre: EDIPUCRS, 1999.
- OLIVEIRA, M. A. *Variation and Change in Brazilian Portuguese: The case of the liquids*. PhD. University of Pennsylvania, 1983.
- OLIVEIRA, M. A. Reanalizando o processo de cancelamento do “r” em final de sílaba. *Revista de Estudos de Linguagem*. v. 6, n. 2, 1997.
- QUEDNAU, L. R. A vocalização variável da lateral. *Letras de Hoje*. Porto Alegre. v. 29, n. 4, p.143-151, dez. 1994.
- RAMOS, A. P. *A generalização estrutural silábica e segmental no tratamento de fala de crianças com desvios fonológicos evolutivos*. In: LAMPRECHT, R. R. (Org.). *Aquisição da Linguagem: questões e análises*. Porto Alegre: EDIPUCRS, 1999.
- RANGEL, G. de A. *Uma análise auto-segmental da fonologia normal: estudo longitudinal de 4 crianças de 1:6 a 3:0*. 1998. Dissertação (Mestrado em Letras) - Instituto de Letras e Artes da PUCRS, Porto Alegre.
- VIDOR, D. C. G. M. Aquisição das líquidas não-laterais por crianças com desvios fonológicos evolutivos – descrição, análise e comparação com desenvolvimento normal. In: 5ª ENAL da PUCRS, 2000, Porto Alegre.
- YAVAS, M., HERNANDORENA, C. L. M.; LAMPRECHT, R. R. *Avaliação fonológica da criança – reeducação e terapia*. Porto Alegre: Artes Médicas, 1992.

Seqüências de (sibilante + consoante) no português de Belo Horizonte

Thaïs Cristófaró-Silva
Universidade Federal de Minas Gerais
KCL

Daniela Mara Lima Oliveira
Universidade Federal de Minas Gerais

Abstract

This paper investigates cases of sound variation involving sequences of (sibilant + consonant) in Brazilian Portuguese (BP). Emphasis is given to the analysis of alveolar (sibilant + stop) sequences as in the word –*festa* [ˈfɛstɐ] – and to the alveopalatal (sibilant + affricate) sequences – as in the word *triste* [ˈtɾiʃtʃi]. The cases of variation we will be dealing with involve the loss of the stop or the affricate where only the sibilant is manifested: *festa* ‘party’ [ˈfɛstɐ] à [ˈfɛsɐ] and *triste* ‘sad’ [ˈtɾiʃtʃi] → [ˈtɾiʃi]. This paper evaluates the role of phonetic properties as potential motivation for the variation: agreement of place of articulation and also agreement of voicing

for the consonants in the sequence is considered. The fact that the sequences involved in the variation are related to a complex syllable pattern – that is, (coda-onset) – will also be explored. Some structural aspects which may condition the phenomena will be addressed: stress placement and the quality of the vowel which precedes the (coda-onset) sequence. It will be suggested that the lack of conditioning structural factors leads us to evaluate this case of variation within the Lexical Diffusion model (WANG, 1969). Lexical Diffusion appears to be the most appropriate approach to analyse the phenomena addressed in this paper, since it accounts for the different behaviour of some structurally similar sequences. Finally, the paper indicates some aspects to be pursued in future research.

0. INTRODUÇÃO

Este trabalho explora casos de variação sonora em seqüências de (sibilante + consoante) no português brasileiro. A ênfase será dada às seqüências de (sibilante + oclusiva) alveolares – como na palavra *festa* [ˈfɛstɐ] – e às seqüências de (sibilante + africada) alveopalatais – como na palavra *triste* [ˈtristʃi]. A variação que pretendemos avaliar relaciona-se aos casos em que a consoante oclusiva ou africada é cancelada e somente uma sibilante ocorre: *festa* [ˈfɛstɐ] → [ˈfɛsɐ] e *triste* [ˈtristʃi] → [ˈtrisi]. Na primeira seção, apresentamos o problema a ser investigado, mostrando que o fenômeno ocorre também em outras línguas. Na segunda seção, avaliamos aspectos fonéticos dos segmentos envolvidos na expectativa de encontrarmos a motivação para tal fenômeno. Na terceira seção, investigamos o papel da estrutura silábica. Na quarta seção, consideramos em detalhes os casos de variação em que uma sibilante ocorre seguida da consoante africada, como em *ginástica*. Nesses casos podemos observar a variação em pelo menos três aspectos: a) uma sibilante alveolar é seguida da africada: [ʒiˈnastʃika], b) uma sibilante alveopalatal é seguida da africada: [ʒiˈnaʃtʃika] e c) uma sibilante alveopalatal ocorre e a africada é cancelada: [ʒiˈnaʃika]. Esses casos de variação serão avaliados em relação ao acento tônico e à vogal precedente. O papel de itens léxicos específicos também é considerado. Os dados são do português de Belo Horizonte, mas o fenômeno em questão foi observado também entre falantes de outras variedades do português brasileiro. Finalmente, indicam-se aspectos a serem abordados em pesquisas futuras.

1. AVALIAÇÃO DO PROBLEMA

Tem sido observado na literatura que pode ocorrer o cancelamento da oclusiva em uma seqüência de (sibilante+oclusiva) alveolares. Em (1) temos exemplos desse fenômeno. Os sons que estamos investigando aparecem em negrito e estão sublinhados nos exemplos que se seguem.

- (1) a. Checo (BARRY & ANDREEVA, 2001)
 de uma distância [vɛvzda:lɔsti] [fvɛzda:losi]
- b. Italiano (BARRY & ANDREEVA, 2001)
 dentista [dentista] [dentisa]
- c. Português brasileiro (CRISTÓFARO-SILVA, 2000)
 festa [fɛsta] [fɛsa]

Os dados em (1) demonstram que, quando uma sibilante alveolar desvozeada ocorre em posição posvocálica e é seguida por uma consoante oclusiva alveolar desvozeada, pode acontecer a simplificação e somente a sibilante ocorre. Ou seja, uma seqüência segmental do tipo [st] pode ocorrer como [s]. Nos exemplos apresentados em (1), temos formas alternantes, isto é, as duas formas ocorrem em competição nessas línguas. Contudo, em uma língua como o inglês, atestamos casos em que a redução de [st] para [s] ocorreu em algumas palavras e, atualmente, não há competição (ocorre somente a forma com a sibilante): *castle* [ka:sɪ] ou *listen* [lɪsn]. Isso implica que o fenômeno em questão pode ter um caráter alternante que, eventualmente, pode se tornar categórico. Por outro lado, seqüências segmentais do tipo [st] ocorrem no inglês atualmente, sem sofrerem a redução para apenas sibilante: *pasta* ‘macarrão’ [pæstə], *vest* ‘corpete’ [vɛst], *street* ‘rua’ [stri:t]. Note que a seqüência segmental em questão [st] comporta-se de maneira diferente em cada grupo de palavras. Nas palavras *castle* e *listen*, a seqüência [st] foi reduzida a [s] e, nas palavras *pasta*, *vest* e *street*, a seqüência [st] foi preservada. Podemos dizer que um determinado fenômeno pode afetar um grupo de palavras em um determinado

estágio (como as palavras *castle*, *listen*), mas não afetar um outro grupo (como as palavras *pasta*, *vest*, *street*). Nós retomaremos esse ponto ao final deste artigo, quando avaliamos o comportamento diferenciado de seqüências estruturalmente semelhantes.

Duas propostas podem ser sugeridas para explicar casos de variação e mudança sonora: neogramática e difusionista. A proposta neogramática assume que as mudanças sonoras são regulares, foneticamente graduais e lexicalmente abruptas. A proposta difusionista assume que as mudanças sonoras não são necessariamente regulares, são foneticamente abruptas e lexicalmente graduais. Este trabalho pretende investigar qual dessas propostas seria mais apropriada para explicar o fenômeno de redução das seqüências de (sibilante + oclusiva) alveolares e de (sibilante + africada) alveopalatais para somente sibilante.

O português de Belo Horizonte apresenta casos semelhantes aos ilustrados em (1). Alguns exemplos são ilustrados em (2) e os sons em questão encontram-se em **negrito** e sublinhados:

(2)	a. festa	['fɛstə]	['fɛsə]
	b. costas	['kɔstəs]	['kɔsəs]
	c. pasta	['pastə]	['pasə]
	d. agosto	[a' gostu]	[a' gosu]

Uma análise dos casos ilustrados em (2) – que envolvem uma seqüência de (sibilante + oclusiva) alveolares – é apresentada em Cristófaros-Silva (2000, 2001b). Um caso adicional apontado em Cristófaros-Silva (em preparação) relaciona-se às seqüências de (sibilante + africada) alveopalatais. Alguns exemplos são apresentados em (3) e os sons em questão encontram-se em **negrito** e sublinhados:

(3)	a. triste	['triʃtʃi]	['triʃi]
	b. vestido	[viʃ' tʃidu]	[vi' ʃidu]
	c. desde	['deʒdʒɪ]	['deʒɪ]
	d. jurisdição	[ʒuriʒdʒi' sãu]	[ʒurizi' sãu]

O fenômeno ilustrado em (3) expressa que, em seqüências de (sibilante + africada) alveopalatais, pode haver variação e somente

a sibilante ocorrer. Esse fenômeno é também observado em limite de palavras. Alguns exemplos são apresentados em (4).

- | | | | |
|-----|-------------------------|-----------------|---------------|
| (4) | a. meus t ios | [meʊʃ 'tʃius] | [meʊ 'ʃius] |
| | b. dois t ipos | ['doiʃ 'tʃipus] | [doi 'ʃipus] |
| | c. Juiz d e Fora | [ʒuiʒdʒi 'fore] | [ʒuiʒi 'fore] |
| | d. giz d e cera | [ʒiʒdʒi 'serə] | [ʒiʒi 'serə] |

Os casos ilustrados em (3,4) ocorrem somente em variedades do português brasileiro em que as oclusivas alveolares são palatalizadas e, neste caso, temos ['tʃia] e ['dʒia], enquanto em variedades que não apresentam a palatalização de oclusivas temos ['tia] e ['dia]. Os casos ilustrados em (3,4) devem, portanto, ser avaliados juntamente com o fenômeno de palatalização. Isso será feito posteriormente. Neste estágio gostaríamos de generalizar afirmando que, em seqüências de (sibilante + oclusiva) alveolares e (sibilante + africada) alveopalatais, pode haver variação e somente a sibilante ocorrer.

Como ocorre com freqüência em casos de variação e mudança sonora, podemos sugerir que a motivação para tal fenômeno seja a semelhança fonética: as consoantes sibilantes e obstruintes compartilham o mesmo ponto de articulação (alveolar ou alveopalatal) e o mesmo grau de vozeamento (vozeado ou desvozeado). Há diferença quanto ao grau e à natureza da estrutura das consoantes em seqüência: uma das consoantes é uma fricativa sibilante e a outra é uma oclusiva ou africada. As consoantes oclusivas e africadas podem ser agrupadas como consoantes obstruintes. Posteriormente, avaliaremos o comportamento das demais consoantes obstruintes do português quando seguidas de sibilantes.

Vale ressaltar ainda que o caso em questão envolve seqüências sonoras que ocorrem em uma estrutura silábica complexa, em que a sibilante aparece na posição posvocálica (de coda) e a oclusiva ou a africada ocorre no início da sílaba na posição de *onset*. Portanto, a complexidade silábica e a semelhança fonética podem ser vistas como fatores favorecedores para que ocorra o cancelamento de uma das consoantes.

Temos dois aspectos a serem investigados. O primeiro deles está relacionado ao comportamento de seqüências de (sibilante + consoante) de maneira geral no português. Pretendemos demonstrar que as seqüências (sibilante + oclusiva) alveolares e (sibilante + africada) alveopalatais comportam-se de maneira distinta de outras seqüências de (sibilante + consoante) em português. O segundo aspecto a ser investigado é a motivação para o fenômeno. Nas próximas páginas, exploraremos o papel desempenhado pela semelhança fonética e pela estrutura silábica. Finalmente, avaliaremos como esse fenômeno está sendo implementado lexicalmente.

2. A SEMELHANÇA FONÉTICA

Vimos na seção anterior que as seqüências de (sibilante + oclusiva) e (sibilante + africada) que são reduzidas a uma sibilante devem concordar quanto ao ponto de articulação e quanto ao grau de vozeamento. Essa observação nos leva a sugerir que há motivação fonética para o fenômeno. Os dois requisitos devem ser satisfeitos: concordância quanto ao ponto de articulação e quanto ao grau de vozeamento. Para corroborar essa hipótese, devemos investigar outras seqüências de (sibilante + consoante) em português. Os exemplos em (5) ilustram casos em que uma sibilante é seguida de uma consoante alveolar no português. Ou seja, as consoantes adjacentes concordam quanto ao ponto de articulação e quanto ao vozeamento.

(5)	asno	['aznu]	° ['azu]	['afnu]
	Islamismo	[izla 'mizmu]	*[iza 'mizmu]	[ifila 'mizmu]

Os asteriscos indicam que somente a ocorrência da sibilante não é possível nesses casos (3ª coluna). Contudo, alternativamente, pode ocorrer o enfraquecimento da sibilante e esta se manifestar como uma fricativa glotal. Esse fenômeno é indicado na coluna mais a direita em (5).

Observe que, nos casos de (5), temos a articulação alveolar das duas consoantes, e ocorre um fenômeno diferente dos fenômenos considerados em (2,3,4). No caso de (5), pode ocorrer o enfraquecimento

do s-posvocálico para [h] e, nos casos de (2,3,4), pode ocorrer o cancelamento da oclusiva ou africada em seqüências de (sibilante + oclusiva) alveolares e (sibilante + africada) alveopalatais. Se o cancelamento da oclusiva ou africada em (2,3,4) fosse motivado exclusivamente pela concordância quanto ao ponto de articulação e quanto ao vozeamento das consoantes envolvidas, esperaríamos que em outras seqüências de consoantes que compartilhem o mesmo ponto de articulação e vozeamento – como em (3,4) – o mesmo fenômeno ocorresse. Esse não é o caso. Podemos concluir que, além da semelhança quanto ao ponto de articulação e vozeamento, é essencial que tenhamos (sibilante + oclusiva) alveolares ou (sibilante + africada) alveopalatais. Os fatos apresentados em (2,3,4,5) mostram também que seqüências estruturalmente semelhantes, ou seja, (sibilante + consoante) podem ter comportamentos diferentes.

Considere os exemplos em (6) que ilustram as demais seqüências de (sibilante + consoante) no português.

(6)	a. áspera	['asperə]	° ['aserə]
	b. esbarro	[iz 'bahʉ]	° [i 'zahʉ]
	c. esfola	[is 'fɔ ə]	° [i 'sɔ ə]
	v. desvio	[dʒiz 'viʉ]	° [dʒi 'ziʉ]
	e. esguia	[iz 'gia]	° [i 'zia]
	f. fiasco	[fi 'askʉ]	° [fi 'asʉ]
	g. asma	['azmə]	° ['azə]
	h. Israel	[isha 'ɛw]	° [isa 'ɛw]

Os dados em (6) mostram que nesses casos não ocorre o cancelamento da consoante que segue a sibilante. Isso é indicado nas formas com asteriscos (coluna da direita). Quando ocorre variação nos exemplos de (6), temos o enfraquecimento da sibilante como uma fricativa glotal: *desvio* [dʒiz 'viʉ] alterna com [dʒifɪ 'viʉ]. Como foi mencionado anteriormente, os casos de enfraquecimento da sibilante em posição posvocálica merecem ainda uma investigação detalhada em português.

Uma outra observação interessante nos exemplos de (6) é que (6a-f) compreendem consoantes obstruintes. As consoantes

[t, d, tʃ, dʒ] também são consoantes obstruintes. Contudo, quando estas últimas consoantes são precedidas de sibilante, o comportamento é diferente dos casos de (sibilante + obstruinte) ilustrados em (6a-f). Em (6a-f) a sibilante permanece ou é enfraquecida para uma fricativa glotal: *desvio* [dʒiz 'viu] ou [dʒiɦ 'viu]. Nos casos em que uma sibilante é seguida de [t, d, tʃ, dʒ], a obstruinte é cancelada e a sibilante permanece: *festa* ['fɛstə] ou ['fɛsə] e *desde* ['dɛʒdʒɪ] ou ['dɛʒɪ]. Novamente, verificamos que seqüências estruturalmente semelhantes, ou seja, (sibilante + consoante) podem ter comportamentos diferentes (compare os dados:2,3,4,5,6).

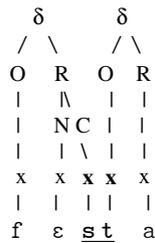
Como generalização, podemos afirmar que, em seqüências do tipo (sibilante + oclusiva) alveolares e (sibilante + africada) alveopalatais, a oclusiva e a africada podem ser reduzidas a uma sibilante. A natureza da estrutura também desempenha um papel importante. As consoantes em seqüência devem ser (sibilante + oclusiva) alveolares ou (sibilante + africada) alveopalatais, observando-se a concordância do lugar de articulação e do vozeamento.

Embora possamos postular que parâmetros fonéticos sejam motivadores para o fenômeno que estamos analisando, devemos observar que, de fato, ocorre o cancelamento de uma consoante (a obstruinte). Certamente, há semelhança fonética na seqüência a ser alterada – sibilante seguida de [t, d, tʃ, dʒ] – mas a semelhança fonética não pode sozinha ser o fator motivador para o cancelamento da consoante oclusiva. Vamos sugerir que a simplificação das seqüências de sibilante seguida de [t, d, tʃ, dʒ] para sibilante procede como decorrência da compressão e redução de esforço articulatorio (BROWMAN & GOLDSTEIN, 1992). Um fato importante que segue a hipótese de compressão e redução de esforço articulatorio é que estruturas silábicas complexas tendem a ser alteradas. Esta hipótese – de compressão e redução de esforço articulatorio em estruturas silábicas complexas – será considerada em detalhes em trabalho futuro (OLIVEIRA, em preparação). O papel da estrutura silábica no fenômeno que estamos investigando será considerado na seção seguinte.

3. A ESTRUTURA SILÁBICA

Esta seção considera o papel da estrutura silábica como fator motivador do fenômeno que estamos analisando. O padrão silábico CV (consoante seguida de vogal) é recorrente em todas as línguas naturais e é, portanto, visto como um padrão silábico universal e não-marcado. Línguas que apresentam um único padrão silábico têm a sílaba CV. Outros padrões silábicos ocorrem em algumas línguas e não em outras e podem existir relações tipológicas entre padrões silábicos: por exemplo, línguas que têm o padrão silábico CCV, como em *prato*, também têm o padrão silábico CVC, como em *larga*. O que nos interessa aqui é que o fenômeno investigado apresenta um padrão silábico marcado em que uma consoante posvocálica é seguida de outra consoante na sílaba seguinte. O diagrama que se segue ilustra a representação da palavra *festa* ['fɛstə] assumido pela Fonologia Autosegmental (GOLDSMITH, 1990).

(7)



Na representação (7), N se refere ao núcleo, O se refere ao onset, R se refere à rima e C se refere à coda. O símbolo δ se refere ao nódulo da sílaba. Um dos desenvolvimentos importantes da Fonologia Autosegmental foi o estabelecimento de uma relação formal entre um *onset* e a coda da sílaba precedente. Nesse contexto ocorrem vários fenômenos fonológicos em várias línguas. Exemplos são a assimilação sistemática de vozeamento ou a perda da consoante nasal em coda e a nasalização da vogal que a precede. Um outro fenômeno que ocorre entre (coda + *onset*) é o cancelamento da consoante em coda. Os exemplos em (8) ilustram casos do português brasileiro em que uma consoante em coda é cancelada:

(8) a. Cancelamento de nasal em coda (MATTOSO CÂMARA, 1970)

campo ['kãpu]

anda ['ãdø]

b. Cancelamento de 'r' em coda (OLIVEIRA, 1981)

porque [puh 'ke] [pu 'ke]

marcha ['mahʃø] ['maʃø]

Contudo, há casos em que a consoante em coda pode permanecer. Os casos em (9) ilustram essa possibilidade. Em (9a) a nasal em coda permanece e a oclusiva que a segue é cancelada. Em (9b) o 'r' em coda permanece e a oclusiva que o segue é cancelada.

(9) a. Cancelamento do 'd' pós-nasal (MOLLICA, 1998)

comendo [ku 'mēdu] [ku 'mēnu]

rindo ['hīdu] ['hīnu]

b. Cancelamento do 'g' precedido por R posvocálico (CRISTÓFARO-SILVA, 2000)

pergunta [pefi 'gūtø] [pe 'hūtø]

perguntei [pefi'gũ 'teɪ̃] [pehũ 'teɪ̃]

Generalizando, podemos afirmar que seqüências consonantais de (coda + *onset*) são estruturas instáveis e podem sofrer alterações. Como estratégia de simplificação de uma estrutura silábica marcada, ou seja, (coda + *onset*) tanto a consoante em coda quanto a consoante em *onset* podem ser canceladas. Para uma melhor compreensão do fenômeno estudado, sugerimos que o cancelamento de consoantes nesse tipo de estrutura silábica complexa, ou seja, (coda + *onset*) pode seguir os padrões silábicos recorrentes na língua em questão. Essa hipótese será investigada em trabalho futuro (OLIVEIRA, em preparação).

Como conclusão da investigação sobre a semelhança fonética e a estrutura silábica, podemos afirmar que esses dois fatores, possivelmente, favorecem e contribuem para a variação atestada nas seqüências sonoras que estamos investigando. Isto porque as seqüências sonoras em questão encontram-se em estruturas silábicas complexas e há semelhança fonética entre os segmentos envolvidos. A hipótese de compressão e redução do esforço articulatório em

estruturas silábicas complexas (BROWMAN & GOLDSTEIN, 1992) será investigada com o objetivo de fornecer uma explicação mais abrangente para o fenômeno em análise.

No restante deste artigo, avaliaremos, em maiores detalhes, os casos em que uma sibilante alveopalatal segue uma africada alveopalatal: *triste* [ˈtɾiʃtʃi]. Restringiremos a análise aos casos em que as duas consoantes são desvozeadas: [ʃtʃ]. Isto porque os casos de seqüências vozeadas do tipo [ʒdʒ] são pouco freqüentes no português (cf. nota 2). Considerando-se o número reduzido de palavras com seqüências do tipo [ʒdʒ], não poderíamos investigar a distribuição estrutural dessas seqüências em relação ao padrão acentual (se a seqüência ocorre em sílaba tônica ou átona) e nem quanto à vogal precedente. Na próxima seção, consideraremos as seqüências de (sibilante + africada) alveopalatais desvozeadas em português.

4. SEQÜÊNCIAS DE (SIBILANTE + AFRICADA) ALVEOPALATAIS

Nesta seção investigamos seqüências segmentais do tipo [ʃtʃ] que passam a se manifestar como [ʃ]. Restringimo-nos aos segmentos desvozeados por razões distribucionais (cf. nota 2). Faz-se pertinente, portanto, considerarmos a distribuição da consoante [ʃ] na estrutura silábica do português.

A sibilante alveopalatal [ʃ] ocorre em todas as variedades do português em posição inicial de sílaba (em início ou meio de palavra). Os exemplos em (10) ilustram estes casos:

(10) Distribuição das sibilantes alveopalatais em início de sílaba:

Vogais orais	a	chave
	ɛ	xeque
	e	chega
	i	faxina
	o	choque
	o	cachorro
	u	chuva

Vogais nasais	ã	chance
	ē	enchente
	ĩ	chimpanzé
	õ	rechonchudo
	ũ	chumbo
Ditongos	oɪ̯	choio
	eɥ̯	encheu
	ãɥ̯	chão

As sibilantes alveopalatais também ocorrem em final de sílaba no português brasileiro e nesse contexto marcam variação dialetal. Na variedade carioca, uma sibilante alveopalatal ocorre, sistematicamente, em final de sílaba. Na variedade paulista, as sibilantes alveolares ocorrem, sistematicamente, em final de sílaba. Já na variedade de Natal, as sibilantes alveopalatais ocorrem seguidas de consoantes alveolares. Finalmente, na variedade mineira de Belo Horizonte, ocorre a sibilante alveolar, exceto quando uma africada alveopalatal ocorre no início da sílaba seguinte. Estes fatos são ilustrados em (11).

(11) Sibilantes em final de sílaba

	Exemplo	Carioca	Paulista	Natal	Mineira
final de palavra	paz	['paʃ]	['pas]	['pas]	['pas]
	giz	['ʒiʃ]	['ʒis]	['ʒis]	['ʒis]
	após	[a 'pɔʃ]	[a 'pɔs]	[a 'pɔs]	[a 'pɔs]
	vezes	['vezɪʃ]	['vezɪs]	['vezɪs]	['vezɪs]
	cusuz	[kuʃ 'kuʃ]	[kus 'kus]	[kus 'kus]	[kus 'kus]
segundo C não alveolar	casca	['kaʃkə]	['kaskə]	['kaskə]	['kaskə]
	lesma	['lezmə]	['lezmə]	['lezmə]	['lezmə]
	vesga	['vezgə]	['vezgə]	['vezgə]	['vezgə]
	espera	[iʃ 'pɛrə]	[is 'pɛrə]	[is 'pɛrə]	[is 'pɛrə]
	esbarra	[iʒ 'baɦə]	[iz 'baɦə]	[iz 'baɦə]	[iz 'baɦə]

	Exemplo	Carioca	Paulista	Natal	Mineira
segundo C alveolar	asno	['aʒnu]	['aznu]	['aʒnu]	['aznu]
	islamismo	[izlã 'mizmu]	[izlã 'mizmu]	[izlã 'mizmu]	[izlã 'mizmu]
	estudo	[iʃ 'tudu]	[is 'tudu]	[iʃ 'tudu]	[is 'tudu]
	festa	['fɛʃtə]	['festə]	['fɛʃtə]	['festə]
	isto	['iʃtu]	['istu]	['iʃtu]	['istu]
segundo C alveopalatal	triste	['triʃtʃɪ]	['tristɪ]	['tristʃɪ]	['triʃtʃɪ]
	haste	['aʃtʃɪ]	['astɪ]	['astʃɪ]	['aʃtʃɪ]
	poste	['pɔʃtʃɪ]	['postɪ]	['postʃɪ]	['pɔʃtʃɪ]
	ginástica	[ʒi 'naʃtʃikə]	[ʒi 'nastikə]	[ʒi 'nastʃikə]	[ʒi 'naʃtʃikə]
	doméstica	[do 'mɛʃtʃikə]	[do 'mɛstikə]	[do 'mɛstʃikə]	[do 'mɛʃtʃikə]

Consideraremos em maiores detalhes os casos em que uma sibilante alveopalatal ocorre na variedade mineira de Belo Horizonte. Mais especificamente, investigaremos os casos em que seqüências do tipo [ʃtʃ] ocorrem. Considere o grupo de exemplos da última coluna (variedade mineira) para o último contexto listado na coluna mais a esquerda (segundo consoante alveopalatal). Essas seqüências estão em competição com outras seqüências alternantes. Isto é ilustrado em (12) para a palavra “ginástica”.

(12) Formas alternantes da palavra *ginástica*

- a. [ʒi 'nastʃikə] seqüência de (sibilante alveolar + africada alveopalatal) e vogal [i]
- b. [ʒi 'naʃtʃikə] seqüência de (sibilante alveopalatal + africada alveopalatal) e vogal [i]
- c. [ʒi 'naʃikə] seqüência de (sibilante alveopalatal) e vogal [i]
- d. [ʒi 'naʃkə] seqüência de (sibilante alveopalatal + outra consoante)

O caso ilustrado em (12d) não será tratado neste artigo. Isto porque esse caso envolve o cancelamento da vogal [i] postônica e merece um tratamento separado, pois implica a criação de uma estrutura silábica complexa: (coda + *onset*). Em relação aos casos (12a-c), observamos que essas três seqüências em competição ocorrem independente de a seqüência ser seguida de sílaba tônica

ou átona. Considere as formas em (13) que são agrupadas quanto à ocorrência da seqüência em relação ao acento tônico. Os exemplos em (13) apresentam alternância entre: [stʃ], [ʃtʃ] e [ʃ].

(13) Distribuição de seqüências do tipo (sibilante + africada) em relação ao acento tônico

Pretônica	Tônica	Postônica	
		Postônica medial	Postônica final
esticadinha	vestido	plástico	triste
destilaria	estilo	ginástica	teste

Os exemplos em (13) mostram que a variação atestada em seqüências de (sibilante + africada) alveopalatais ocorre em sílaba tônica ou átona (pretônica e postônica). Podemos afirmar que não se aplica o condicionamento estrutural da tonicidade ao fenômeno. Ou seja, o fenômeno não parece ser condicionado pela posição da seqüência de (sibilante + africada) em relação à tonicidade. Podemos afirmar também que não há condicionamento estrutural quanto à vogal que precede essa seqüência. A consoante africada é sistematicamente seguida por [i] e não há restrição quanto à vogal que precede a sibilante posvocálica, como indicado nos exemplos que se seguem. Em (14) ilustramos casos em que qualquer uma das vogais orais do português precede a sibilante posvocálica.

(14) Vogal + [ʃtʃ]

a	ginástica
ɛ	veste
e	investigado
í	triste
o	poste
o	rostinho
u	rústico

Se não há condicionamento estrutural que regule a redução de (sibilante + africada) alveopalatal a uma sibilante, devemos buscar elementos adicionais para que possamos compreender melhor o

fenômeno. A nossa proposta é a de que esse fenômeno seja investigado na perspectiva da teoria da Difusão Lexical (WANG, 1969). Esse pretende ser o foco central de pesquisa futura (OLIVEIRA, em preparação). Dentro dessa abordagem, investigaremos como a variação está sendo implementada no léxico, ou seja, palavra por palavra. Como suporte adicional a essa proposta, consideraremos aspectos extralingüísticos como: idade, sexo, grau de instrução e estilo de fala. Pretendemos avaliar, ainda, o comportamento de falantes individualmente (OLIVEIRA, 1992).

A hipótese da Difusão Lexical nos parece pertinente por permitir expressar que as mudanças afetam palavras em um tempo determinado. Note que os pronomes ilustrados em (15) apresentam no português atual quase que categoricamente a forma em que somente a sibilante ocorre (sendo que a oclusiva ou a africada é cancelada). Esses pronomes são certamente muito freqüentes. A freqüência de ocorrência é possivelmente um fator que pode contribuir para a implementação de uma mudança. Contudo, o que neste momento nos interessa é que os itens em (15) se fundiram em apenas uma forma fonética em que somente a sibilante ocorre (originalmente estes exemplos tinham forma e significados diferentes).

(15) Palavras que fundiram forma e significado

a. este	esse	[¹ esɪ]
b. esta	essa	[¹ ɛsə]
c. isto	isso	[¹ isʊ]
d. neste	nesse	[¹ nesɪ]
e. nesta	nessa	[¹ nesə]
f. nisto	nisso	[¹ nisʊ]
g. deste	desse	[¹ desɪ]
h. desta	dessa	[¹ dɛsə]
i. disto	disso	[¹ dʒisʊ]

Obviamente, as formas ilustradas em (15) podem apresentar a seqüência [st] em estilo formal, sobretudo em casos de leitura. Na fala coloquial e tipicamente em estilo informal, geralmente ocorre apenas

a forma com a sibilante [s], como nos exemplos ilustrados em (15). Poderíamos afirmar ainda que a grande maioria dos falantes não distingue diferentes significados para os pares: *este/esse*, *neste/nesse* etc.

Consideremos um aspecto interessante que se refere aos itens (15a,d,g): *este*, *neste*, *deste*. Note que, se tais itens tivessem sido submetidos à mudança no estágio atual, esperaríamos ter as seguintes realizações fonéticas: ['eʃɪ, 'neʃɪ, 'deʃɪ], respectivamente para *este*, *neste*, *deste*. Compare, por exemplo, o comportamento da palavra *leste*. Em alguns dialetos que não apresentam a palatalização de oclusivas alveolares, essa palavra é pronunciada como *leʃtʃe* (com seqüência de (sibilante + oclusiva alveolar)). Em dialetos que apresentam a palatalização de oclusivas alveolares, temos as pronúncias: *leʃtʃje*, *leʃtʃje* ou *leʃje* (cf. (12)). Na pronúncia *leʃtʃje*, a palatalização não afeta a sibilante em posição posvocálica. Na pronúncia *leʃtʃje*, a sibilante posvocálica é palatalizada e ocorre como [ʃ]. Finalmente, na pronúncia *leʃje*, a africada é cancelada e somente a sibilante ocorre.

Se compararmos o comportamento da palavra *leste* com o comportamento do pronome *este* (e também *neste*, *deste*), esperaríamos que *este* apresentasse as seguintes pronúncias: *eʃtʃje*, *eʃtʃje* e *eʃje*. De fato, em estilo formal ou de leitura, podemos atestar as formas *eʃtʃje* e *eʃtʃje*. Contudo, a forma esperada *eʃje* não ocorre e, de fato, temos a pronúncia *eʃje* para *este*. Observamos que formas estruturalmente semelhantes – *leste* e *este* – apresentam comportamentos diferentes quanto aos casos de variação em que somente a sibilante ocorre. A forma *leste* apresenta a pronúncia *leʃje* (mas não *leʃje*) e a forma *este* apresenta a pronúncia *eʃje* (mas não *eʃje*). A pronúncia diferente para seqüências estruturalmente semelhantes demonstra que o grupo de palavras em (15) deve ter sido afetado em um momento no qual a palatalização não era operante. Se a mudança estivesse relacionada ao processo de palatalização, deveríamos atestar formas que realmente não ocorrem no português contemporâneo (*eʃje*, *neʃje*, *deʃje*).

Por outro lado, em todos os dados em (15), observamos um comportamento análogo às seqüências de (sibilantes + oclusivas)

alveolares. Veja que, na palavra *feira*, que não se relaciona ao processo de palatalização de oclusivas alveolares, temos as pronúncias *fe[st]a* ou *fe[s]a*. Note que, tanto em *este* quanto em *feira*, uma (sibilante alveolar) ocorre: *fe[s]a*, *e[s]e*. Esse comportamento aparentemente inadequado pode ser facilmente capturado pela teoria da Difusão Lexical. Isto porque nesse modelo as mudanças sonoras são implementadas no léxico gradualmente. Sendo assim, a mudança sonora observada nos casos como os de (15) deve ter sido implementada numa fase diferente das palavras que têm a mesma mudança implementada neste momento (neste estágio temporal do português de Belo Horizonte, a mesma mudança seria o cancelamento da oclusiva ou africada em seqüências de (sibilante + oclusiva ou africada) sendo que somente a sibilante ocorre).

5. CONCLUSÃO

Este artigo teve por objetivo avaliar casos de variação sonora em seqüências de (sibilante + consoante) no português brasileiro. A ênfase foi dada às seqüências de (sibilante + oclusiva) alveolares – como na palavra *feira* [ˈfɛstɐ] – e às seqüências de (sibilante + africada) alveopalatais – como na palavra *triste* [ˈtɾiʃtɕi]. A variação que avaliamos relaciona-se aos casos em que a consoante oclusiva ou africada é cancelada e somente a sibilante ocorre: [ˈfɛsɐ] e [ˈtɾiʃi].

Vimos nas seções precedentes que a motivação fonética parece contribuir para a ocorrência do fenômeno em questão, pois as consoantes envolvidas devem ter o mesmo ponto de articulação e o mesmo grau de vozeamento. Há ainda a importância de a estrutura silábica ser complexa. A hipótese básica a ser investigada em um trabalho futuro é a de que a variação envolvendo (sibilante + oclusiva) alveolares e (sibilante + africada) alveopalatais opera devido a compressão e redução de esforço articulatorio, sendo que estruturas silábicas complexas tendem a ser alteradas. Essa hipótese investigará a motivação dos casos de variação que estamos analisando.

O modelo da Difusão Lexical (WANG, 1969) permite expressar a falta de condicionamento estrutural do fenômeno. Ou seja, o fato de que nem o padrão acentual e nem a natureza da vogal que precede a seqüência segmental a ser alterada condicionam o fenômeno (cf. exemplos 13, 14). A proposta da Difusão Lexical permite ainda explicarmos o comportamento distinto de seqüências segmentais estruturalmente semelhantes (cf. 15). Em uma pesquisa futura, pretendemos explorar como o modelo da Difusão Lexical permite-nos explicar a implementação da mudança que estamos analisando.

A avaliação do fenômeno envolvendo seqüências de (sibilante + consoante) no português brasileiro contemporâneo apresentada neste trabalho oferece pistas iniciais de análise e indica caminhos propícios a serem explorados. Esperamos que uma visão mais completa do fenômeno possa ser oferecida. Este artigo indica aspectos a serem explorados em pesquisa futura com o objetivo de obtermos uma análise mais acurada do fenômeno em questão (OLIVEIRA, em preparação). Esperamos também que certos aspectos ainda não explorados na literatura sejam investigados. Dentre eles, destacamos os casos de aspiração de s-posvocálico (cf. (5,6)) e o comportamento de seqüências de (sibilantes + consoantes) em dialetos que não apresentam palatalização de oclusivas alveolares (ver nota 13). Seria relevante também investigar por que a consoante oclusiva ou africada é cancelada e não a sibilante. Vimos em (8,9) que tanto a consoante que ocorre na posição de coda, quanto a consoante que segue a posição de coda podem ser canceladas. Seria importante investigar parâmetros tipológicos em relação a qual das consoantes da seqüência (coda + *onset*) pode ser cancelada e idealmente definir as condições que favorecem ou bloqueiam o cancelamento da consoante em questão.

NOTAS

¹ Os dados são do projeto intitulado “Difusão lexical: estudo de casos do português brasileiro” (CNPq: 20.2424-86.3 e 301029-89.0).

² Seqüências de (sibilante + oclusiva) vozeadas [zd] e [ʒdʒ] são pouco freqüentes no português. Em uma pesquisa ao Dicionário Michaelis www.uol.com.br/michaelis/, verificamos que ocorrem somente 22 palavras com a seqüência sonora [zdʒ] e [ʒdʒ], em uma listagem de aproximadamente 200 mil palavras. Já as seqüências [stʃ] e [ʃtʃ] aparecem em 2.453 palavras (dentre 200 mil). Neste artigo, restringimos a análise a estes últimos casos.

³ As consoantes obstruintes compreendem as categorias de oclusivas, fricativas e africadas.

⁴ O fenômeno de enfraquecimento da sibilante em coda é verificado no português (cf. GRYNER & MACEDO, 1981; PESSOA, 1986; RONCARATI, 1988; SCHERRE & MACEDO, 1989 *apud* AULLER, 1992) e em espanhol (BYBEE, 2001). Esse tópico ainda merece uma discussão mais exaustiva no português brasileiro.

⁵ O caso de (7a) ocorre sistematicamente nas formas de gerúndio, mas pode ser observada também em outras categorias como advérbios e substantivos (MOLLICA, 1998; CRISTÓFARO-SILVA, 2000). O caso de (7b) parece ser lexicalizado. Ou seja, [g] é cancelado em *pergunta*, *perguntado*, *perguntei* etc., mas não em outras palavras com seqüências sonoras semelhantes: *argumento*, *erguer*; etc. Este caso merece ainda uma descrição mais detalhada.

⁶ Sem entrarmos em detalhes, o que argumentamos aqui é que uma palavra como *pergunta* e semelhantes ocorre como [pehũtə] porque a seqüência sonora [ehũ] é mais recorrente do que [efigũ]. Por outro lado, em palavras como *argumento*, *erguer*, etc., a seqüência [ahũ] e [ehel] é menos recorrente do que [ahĩg] e [efĩg].

⁷ A indicação das variedades dialetais tem caráter de agrupar variedades com os traços descritos: distribuição das sibilantes posvocálicas. Na variedade paulista, nos referimos às comunidades que não apresentam a propriedade de palatalização de oclusivas alveolares.

⁸ Excluimos as vogais nasais e ditongos dos exemplos em (14) por serem menos recorrentes em português do que as vogais orais. Contudo, note que há exemplos que ilustram a ocorrência de vogais nasais seguidas da seqüência [ʃtʃ]: ‘in[ʃtʃ]lituição’ ou ‘con[ʃtʃ]lituição’. Ditongos decrescentes raramente ocorrem seguidos da seqüência [ʃtʃ]. Tipicamente, há restrições quanto à ocorrência de ditongos decrescentes seguidos de consoantes posvocálicas, como em *cáustica*, *exaustivo*, etc.

⁹ No *corpus* do projeto “Difusão Lexical: estudo de casos do português brasileiro” que conta com 47.883 palavras, na fala espontânea, foram constatadas apenas cinco ocorrências do pronome *isto* e uma ocorrência do pronome *disto*. Essas formas foram pronunciadas por membros da mesma família. Os demais pronomes demonstrativos apareceram apenas com a forma em que somente a sibilante alveolar [s] ocorre.

¹⁰ Seria interessante avaliar se em dialetos que não-palatalizam as oclusivas alveolares a pronúncia *lɛs|e* ocorre. Não possuímos dados neste estágio da pesquisa quanto a essa pronúncia

REFERÊNCIAS BIBLIOGRÁFICAS

AULLER, M. A difusão lexical de um fenômeno de aspiração no português. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 1, ano 1, 1992.

BARRY, W.; ANDREVA, B. Cross-languages similarities and differences in spontaneous speech patterns. *Journal of the International Phonetic Association*, v. 31, n.1, p. 51-66. CUP, 2001.

BROWMAN, C.; GOLDSTEIN, L. Articulatory Phonology: An overview. *Phonetica* 49, p.155-180, 1992.

BYBEE, J. *Frequency and the emergence of linguistic structure*. Amsterdam: Philadelphia, PA: John Bejamins Publish, 2001.

CRISTÓFARO-SILVA, T. Sobre a Quebra de Encontros Consonantais no Português Brasileiro. *Estudos Lingüísticos*. São Paulo, v. 29, p 522-527, 2000.

CRISTÓFARO-SILVA, T. *Fonética e fonologia do português*: roteiro de estudos e guia de exercícios. São Paulo: Contexto. 2001a.

CRISTÓFARO-SILVA, T. Difusão Lexical: Estudo de casos do português brasileiro. In: MENDES, E. A. de M., OLIVEIRA, P. M.; BENN-IBLER, V. (Org.). *O novo milênio: interfaces lingüísticas e literárias*. Belo Horizonte: UFMG/FALE. p. 209-218, 2001b.

CRISTÓFARO-SILVA, T. Difusão lexical: estudo de casos do português brasileiro. Relatório final de Pesquisa. (em preparação).

GOLDSMITH, J. *Autosegmental and Metrical Phonology*. Oxford: Basil & Blackwell, 1990.

GRYNER, H.; MACEDO, A. La prononciation du s post-vocalique deux processus de changement linguistique en portugais. In: *Variation omnibus*. Canada, Linguistique Research, p. 135-140, 1981.

- MATTOSO CÂMARA JR., J. *Estrutura da língua portuguesa*. 14. ed. Petrópolis, RJ: Vozes, 1970.
- MICHAELIS - Moderno Dicionário da Língua Portuguesa. Edições Melhoramentos. Disponível em: www.uol.com.br/michaelis/, 2002.
- MOLLICA, M. C. *Influência da fala na alfabetização*. Rio de Janeiro: Tempo Brasileiro, 1998.
- OLIVEIRA, M. A. Reanálise de um problema de variação. In: *Português: Estudos lingüísticos*, Série Estudos, v.7, FISTA, 23-51, 1981.
- OLIVEIRA, M. A. Aspectos da difusão lexical. *Revista de Estudos da Linguagem*, Belo Horizonte, v. 1, p. 31-41, 1992.
- OLIVEIRA, M. A. Reanalizando o processo de cancelamento do (r) em final de sílaba. *Revista de Estudos da Linguagem*. Belo Horizonte, v. 6, n. 2, 1999.
- OLIVEIRA, Daniela Mara Lima. Seqüências de (sibilante +africada alveopalatal) no português de Belo Horizonte. Projeto de mestrado (em preparação).
- PESSOA, M. A. F. C. Um processo de enfraquecimento na fonologia do português. Monografia apresentada no curso de Doutorado em Lingüística da Faculdade de Letras da UFRJ, 1986.
- RONCARATI, C. N. S. Enfraquecimento das fricativas sonoras. In: Relatório final à FINEP – Projeto Dialetos Sociais Cearenses. Fortaleza Universidade Federal do Ceará. Convênio FINEP; CCPC 41.85.0655.00, 1988.
- SCHERRE, M. M. P.; MACEDO, A. T. Variação e mudança: o caso da pronúncia do S-posvocálico. In: Relatório final à FINEP – Projeto: Mecanismos Funcionais do uso Lingüístico. Rio de Janeiro. UFRJ. Convênio FINEP-UFRJ, 1989.
- WANG, W. S-Y. Competing changes as a cause of residue. *Language*, n. 45, 1969.

O alçamento das vogais médias pretônicas e o conceito de léxico com armazenamento exemplar

Maria do Carmo Viegas
Universidade Federal de Minas Gerais

Abstract

The main purpose of this paper is to make evident that many lexical items of the portuguese language cannot be seen as exceptions to the pretonic mid vowels raising rules as they have been considered, and that lexical diffusion is not lexical analogy.

Este trabalho se baseia na análise dos dados levantados em duas áreas da região metropolitana de Belo Horizonte diferenciadas socioeconomicamente (VIEGAS, 1987) e na análise histórica, feita em Viegas (2001), de diversos itens lexicais em textos portugueses dos séculos XII a XIV, em textos portugueses do século XVI¹ e em dicionários (BLUTEAU, 1712-1721; COROMINAS, 1954-1957; CUNHA, 1982; FERREIRA, 1975).

Em Viegas (1987), além de uma proposta de diferentes regras de alçamento, ou elevação do traço de altura, das vogais médias pretônicas (minino, cunversa etc.), uma regra para o /e/ – harmonização vocálica favorecida pela vogal alta seguinte – e outra para o /o/ – redução favorecida pelas consoantes adjacentes –, mostro que há evidências de uma atuação lexical relacionada ao alçamento. Encontrei essa relação evidenciada em formas como *Peru* e *piru*, entre outras. Nesse par, temos, hoje, itens com significados diferentes, um item alçado e outro não, mas que tiveram origem no mesmo item (*piru* veio do topônimo *Peru*, segundo CUNHA, 1982). Se observarmos alguns itens, veremos que o item alçado é aquele cujo uso normalmente é feito em situações mais familiares (menos marcadas) ou em situações menos prestigiadas (*piru* também é considerado um item chulo). Assim, nas formas em que a vogal média está em oposição distintiva em relação à vogal alta, aqueles itens que têm um uso considerado menos prestigiado ou aqueles itens mais familiares alçam (*purção*, com o significado de “muita quantidade”, traz o rótulo brasileiro) e aqueles cujo uso é socialmente considerado mais prestigiado ou mais marcado com o prestígio não alçam (*porção* é usado, especialmente, em restaurantes, ou ambientes marcados (MILROY, L., 1987), e *Peru* é inclusive escrito com letra maiúscula).

Viegas (1987) é um trabalho que reconhece a implementação das regras² como um processo lexicalmente gradual. No entanto, maiores estudos fizeram-se necessários para entendermos melhor o processamento do alçamento, já que foram propostas regras de descrição do processo, mas com implementação lexicalmente gradual.

O primeiro passo para melhor entendermos o alçamento foi analisarmos a história de alguns itens. Nessa análise ficou constatado que alguns itens considerados como atingidos pelo processo denominado alçamento talvez não deveriam estar aí incluídos (VIEGAS, 2001). Os itens que chegaram até nós por meio de outras línguas que não o latim podem, muitas vezes, ter vindo com a vogal já alta. Assim, *tupete*, *ciroulas*, *supapo*, *tumate*, *algudão*, *almufada*, *bulinar*, *buneca*, *bunito*, *butão*, *mulambo*, *muleke*, *muqueca* e outros foram separados da amostragem dessa análise. Essa é uma possibilidade que observamos hoje com os itens *taliban* (muitas vezes escrito *taleban*) e *i-mail* (escrito *e-mail*). Nesses casos, a grafia de e e o não corresponde geralmente à pronúncia de vogal média e não há motivo para se falar em alçamento.

Separei também os itens que vieram de uma vogal alta latina e cuja pronúncia com vogal alta foi detectada ou depreendida muito cedo no português:

murcego < do a. port. mur (séc. XIII), rato, + cego.

assubiar < séc. XIV, assuuyar (do lat. adsibilare)

bizerro < do lat. hisp. *ibicerra, *ibicirra

gingiva < do lat. gingiua

governo < do lat. vulgar guuernare

mosquito < do lat. musca

truvão < do lat. vulg. turbo, - onis

pequeno < do lat. vulg. pitinnu - séc. XIII

assoc. a uma base expressiva pikk = 'pequenez'

etc.

Esses itens também devem ser analisados separadamente, assim como os prováveis nivelamentos analógicos (*milhor* e *pior*;

simestre e bimestre). Nesses itens, caso haja variação, creio que devemos falar em “medialização”.

Obviamente temos um problema metodológico se depreendermos sons de formas escritas na análise. Assim, se temos a vogal alta na escrita em determinada época, pressupomos que, pelo menos em uma variante, tivemos a pronúncia de vogal alta; já se temos a vogal média na escrita em determinada época, podíamos, ou não, ter a pronúncia de vogal alta em pelo menos uma variante, ou seja, o grau de incerteza é maior.

Tomo, então, a escrita como reveladora da pronúncia, pelo menos no período chamado de ortografia fonética (dos primeiros documentos até o século XVI). O certo é que, se existe um período ortográfico que pode nos dar pistas sobre a pronúncia das palavras, este é o chamado período fonético, posto que as convenções ortográficas não estavam bem estabelecidas.

Segundo Naro (1973: 28):

Em suma, há duas tendências evidentes. Até Barreto (1671) o sistema ortográfico não apresentava nenhum problema pois não havia uma separação notável entre a ortografia e o uso da língua falada. Até o tempo de Feijó (1739), a confusão era tão grande que medidas normativas se fizeram necessárias. Mais ou menos na metade entre os dois extremos opostos (1696), a máxima de Quintiliano [‘sic scribendum quomodo sonat.’] foi apropriadamente suprimida e até recentemente não se permitiu que ressurgisse.

A análise histórica que ora faço, depreendendo sons da grafia, combinada com as análises feitas por estudiosos da história da língua e com a análise dos sons tal qual eles se apresentam hoje forma um quadro que se sustenta mais firmemente.

A interpretação que dou neste trabalho é uma versão com um menor número de “exceções”, pois estou separando os empréstimos e os itens que, provavelmente, já foram incorporados com a vogal alta. À primeira vista, os dados talvez se encaixassem numa análise do tipo feito em Kiparsky (1995), mas, diferentemente desse autor, encontro subsídios para afirmar que difusão lexical não é analogia lexical, conforme veremos a seguir.

Vamos então à análise histórica dos itens:

e: Itens alçados (sem os empréstimos) – século XIII.³

Empréstimo: origem não latina

- | | |
|---------------------------------|-------------------------------------|
| 1) aparecia, aparecimento (XV) | 2) apelido, apelidar (XIV) |
| 3) aprendi | 4) arcebispo (l. ecles.) |
| 5) cemitério, cymiteiro | 6) conhecia, connocer (XIII) |
| 7) consentir, consintidor (XIV) | 8) dependurar, dependorar, pendurar |
| 9) derrubar, derrubar | 10) devia |
| 11) ferir | 12) fervido |
| 13) gengibre, gingibre (XV) | 14) jejum, jajúm |
| 15) medida, midida | 16) mentir |
| 17) messias | 18) nenhúm, nengun, nigu (XIV) |
| 19) oferecido | 20) pedir, pitiçon (XIII) |
| 21) pendurar, pendorar | 22) penitência |
| 23) percebido | 24) perigo |
| 25) pertencia | 26) peru (do topônimo) |
| 27) preguiça, priguíça (XIV) | 28) prendi |
| 29) recebi, recybymento | 30) revelia |
| 31) seguir | 32) segundo |
| 33) segurar | 34) sentir |
| 35) servir | 36) testemunho, testimoyo |
| 37) tremido | 38) vencido |
| 39) vendido | 40) vestir |

e: Itens alçados (sem empréstimos) – século XIV:

- | | |
|----------------------------|----------------|
| 1) abreviar, abriviar (XV) | 2) agradescido |
| 3) bexiga (vexiga) | 4) defunto |
| 5) desistir | 6) feliz |
| 7) perseguir, persiguir | 8) repetir |
| 9) verruga (berruga) | |

e: Itens alçados (sem empréstimos) – século XV:

- | | |
|------------------------------|---------------------|
| 1) arripiar, arripiado (XVI) | 2) digerir, digirir |
| 3) impedir | 4) intestino |
| 5) mexido | 6) veludo |

e: Itens alçados (sem empréstimos) – século XVI:

- | | |
|--------------|---------------|
| 1) acreditar | 2) cerzir |
| 3) conseguir | 4) inferir |
| 5) mexerica | 6) ofendido |
| 7) precisa | 8) pressentir |
| 9) querido | 10) seringa |

e: itens alçados (sem empréstimos) – século XVII:

- | | |
|--------------|-------------|
| 1) aperitivo | 2) preferir |
| 3) relinchar | |

e: Itens alçados (sem empréstimos) – século XVIII:

- | | |
|--------------|------------|
| 1) exigência | 2) existir |
| 3) peludo | 4) pepino |

e: itens alçados (sem empréstimos) – século XIX:

- | | |
|-------------|-------------|
| 1) agredir | 2) apreciar |
| 3) beliscar | 4) besuntar |
| 1) decidir | 6) existir |
| 7) impetigo | 8) penico |

Observando os dados, podemos concluir que o ambiente vogal alta seguinte foi bastante favorecedor ao alçamento do item. Passemos, pois, ao estudo das listas dos itens não alçados e que possuíam ambiente de vogal alta seguinte.

e: Itens com ambiente de vogal alta seguinte, não alçados e separados por séculos (sem os empréstimos)
Empréstimo: origem não latina

Séc. XIII

- 1) perdiz
- 2) petição, piticon

Séc. XIV

- 1) cerviz

Séc. XV

- 1) cenáculo

Séc. XVI

- | | |
|--------------|---------------|
| 1) averiguar | 2) cerúleo |
| 3) crepitar | 4) crepúsculo |
| 5) crespidão | 6) perícia |
| 7) Netuno | |

Séc. XVII

- 1) assentir
- 2) atenuar
- 3) cernir
- 4) clandestino

Séc. XVIII

(não há registro na amostragem)

Séc. XIX

- 1) celícola

Séc. XX

(não há registro na amostragem)

Observamos que o número de itens não alçados e com ambiente para tal não é grande, indicando que o processo de harmonização vocálica não é tão “excepcional” como afirma Teyssier (1997: 101).⁴ O processo só não foi/é uniforme, ou abrupto em termos lexicais, foi/é lexicalmente gradual.

Os itens que chamamos populares, do dia-a-dia, são os primeiros itens a serem alçados, indicando que os primeiros itens a se submeterem a processos como o alçamento são, na maioria, os itens adquiridos no meio familiar, em oposição aos aprendidos e objeto de instrução.⁵

Assim, ao que parece, a vogal e passou a i, através de um processo inicial de harmonização vocálica, atingindo o léxico item a item, ou grupos de itens a grupos de itens; posteriormente, então, houve a redução do i, em Portugal.

O processo de harmonia vocálica não é excepcional, ele é *lexical*. Os itens não alçados são, em geral, mais formais, mais especializados, mais eruditos.

É interessante notarmos que temos itens, com o primeiro registro nos séculos XIII a XVII, que possuem ambiente para o alçamento (*perdiz, cerviz, crepúsculo, Netuno*, etc.) e que não foram registrados alçados, enquanto outros, nessas condições descritas, foram registrados alçados (*cimitério, medida, ninbum*, etc.), mostrando-nos a atuação lexical.

Assim como foi feito para o e, separamos os empréstimos e avaliamos a possibilidade de muitos itens terem o o somente na grafia e há muito terem a pronúncia [u] em pelo menos uma variedade do português, que acabou constituindo a variedade por nós estudada. Uma possibilidade explicativa da análise é que alguns itens que vieram de vogal alta (*gurdura, murcego, mosquito*, etc.) e que tiveram um registro com vogal alta no português do Brasil recém descoberto têm a pronúncia de vogal alta hoje com base na pronúncia daquela época. Mas, no caso do o, a elevação da vogal média, tendo em vista o ambiente de harmonização vocálica, precisa ser considerada anteriormente, no latim, ou na sua passagem para o português, para que tenhamos maior sistematicidade. Mesmo assim, o número de exceções a uma possível regra de harmonização vocálica é muito maior do que no caso do e. Teremos que nos esforçar muito mais se quisermos diminuir o número de exceções e, ainda assim, o número de exceções é grande.

Temos alçados *budega, chucalho, cumer, cumpadre, cunsertar, custela, custeleta, encumendar e tulerar*. E os itens *culher* (s.) e *colher* (v.)? Mais uma vez temos evidências de que estamos diante de um processo de difusão lexical, pois o processo atinge alguns itens e não atinge outros.

A extrapolação do ambiente vogal alta seguinte (*budega, bucejar, apudrecer*, etc.) é evidente. Então poderíamos pensar que o ambiente favorecedor talvez fosse outro. Qual? Consoantes adjacentes? Temos vários itens com a mesma consoante precedente, alçados e não-alçados (*cunsertar, cumer, custela, coligir, comício, contusão*, etc.).

Se compararmos alguns itens com o primeiro registro no século XIII e que não alçam (*apostila, coluna, condição, confirmar*,

consentir, etc.) com outros alçados e com primeiro registro no século XIII (*acumpanbar*, *acustumar*, *adurmecer*, *apudrecer*, *apruveitar*, *furmiga*, etc.), observamos que, de modo geral, são os itens passados de geração a geração no meio familiar os primeiros a serem alçados.

Há uma extrapolação do ambiente considerado favorecedor ao processo e há itens com ambiente favorecedor que não alçam, conforme analisamos. Isto é verdadeiro particularmente no caso do /o/. Ao que parece, o /o/ teve o seu processo extrapolando os ambientes favorecedores antes do /e/. Então a transformação do [o] em [u] foi atingindo o léxico gradualmente, até se completar em Portugal, mas não no Brasil.

Por volta de 1800, a transformação [o] em [u] está consumada. (TEYSSIER, 1997:75)

Como seria possível explicar as exceções hoje e, principalmente, nos primeiros séculos do processo, se não houvesse aí uma questão lexical atuando? Alguém poderia dizer: aqueles itens foram todos alçados (*perdiz*, *crepúsculo*, *apostila*, *covil*, etc.), todos atingidos pelo processo, mas a grafia não revela isto. Essa afirmação é difícil de ser sustentada pois:

- 1º) – os estudos da língua não mostram que o processo tenha atingido todo o léxico inicialmente; ao contrário, a variação existente na escrita indica que talvez o processo tenha sido lexicalmente gradual (pela grafia de vogal alta em *alguns* itens,⁶ mas não em todos) e também foneticamente gradual (pela confusão da grafia, ora alta ora média, em um mesmo item, no mesmo texto);⁷
- 2º) – o português do Brasil – tomado como uma etapa do português de Portugal – é mais um indício de que o léxico não foi atingido abruptamente, pois aqui, como já foi mencionado, alguns itens alçam, outros não.

Como pudemos observar, os primeiros itens a se submeterem ao processo foram os populares do “patrimônio hereditário” (terminologia de Teyssier). O desprestígio dos portugueses no Brasil,

assim como o desprestígio do alçamento, era evidente como nos disse Naro (1973).⁸ Ao passo que as palavras eruditas (objeto de instrução) constituem valores, prestígio, e desfavorecem o processo de alçamento. Assim, a lista do alçamento decresce com o tempo, com o aumento dos empréstimos e com o fortalecimento da escrita. A lista do não-alçamento cresce com o tempo. Uma possibilidade interpretativa é que a mistura de dialetos acirra o estigma e desperta um certo grau de consciência do estigma do alçamento, que estava vinculado aos grupos do início da colonização, e que se perpetuou até os dias de hoje. Identifico uma segunda etapa do processo de alçamento, produtiva ainda hoje, que está vinculada a um uso jocoso e pejorativo (piscucinho, litrinha etc.).

É interessante observarmos que estamos fazendo uma análise da região de Belo Horizonte e uma possibilidade explicativa é que o alçamento não tenha acontecido dessa maneira em outras regiões em que não houve uma marcação de estigma para o alçamento, o qual, então, propagou-se para outros itens, inclusive empréstimos. Em algumas regiões, o alçamento pode ter atingido muitos itens do léxico e em outras não tantos, dependendo da formação dos grupos sociais, da época da formação e do grau de estigma e marcação que os grupos que se estabeleceram nessas regiões tenham atribuído ao alçamento.

Podemos ver que, mesmo nesta análise em que separo os estrangeirismos e os itens que já foram incorporados ao léxico português com a vogal alta, evidencia-se a *implementação lexical*. Isto não implica dizer que não houve ou não há ambientes favorecedores.

Labov (1994) se utiliza da teoria da Fonologia Lexical (FL) para a explicação dos processos de mudança. Teoria essa (FL) que postula dois tipos de regras, lexicais e pós-lexicais. O alçamento (“*vowel shifts in place of articulation*”) estaria enquadrado como um processo de mudança sonora regular (oposto à difusão lexical) e, segundo Labov, seria descrito por uma regra pós-lexical.

No caso do alçamento de vogais médias pretônicas no dialeto da região de Belo Horizonte, hoje, se formos estabelecer regras, essas

parecem ter exceções (*murango, simestre*), parecem ser sensíveis à fronteira de morfema (*renascimento, co-autor*), parecem ser cíclicas (*cunhicia, piquinininho*), ou seja, têm características das regras lexicais.

Tivemos um processo de harmonização vocálica (regra lexical, segundo LEE, 1995: 2), nitidamente marcado no caso do /e/, e um processo de neutralização (regra pós-lexical, conforme LEE, 1995: 2), que se desenvolveu notadamente em Portugal. No caso do /o/, o processo de neutralização atingiu maior número de itens do que no caso do /e/, na região de B.H. Hoje temos, ainda, um processo produtivo associado ao uso pejorativo do diminutivo –inho.

Argumenta Labov (1994:542) que:

Regular sound change is the result of a gradual transformation of a single phonetic feature of a phoneme in a continuous phonetic space. It is characteristic of the initial stages of a change that develops within a linguistic system, without lexical or grammatical conditioning or any degree of social awareness ('change from below').

Lexical diffusion is the result of the abrupt substitution of one phoneme for another in words that contain that phoneme. The older and newer forms of the word will usually differ by several phonetic features. This process is most characteristic of the late stages of an internal change that has been differentiated by lexical and grammatical conditioning or has developed a high degree of social awareness or of borrowings from other systems ('change from above'). (Grifo meu)

Diferentemente de Labov, penso que o modelo neogramático e o difusionista não diferem na questão da regularidade. Eles diferem na maneira como as exceções são tratadas nos dois modelos (WANG e LIEN, 1993: 348).

Labov (1994) coloca as regras lexicais como foneticamente abruptas – a forma nova e a forma antiga diferindo em vários traços fonéticos – numa fase tardia de um processo. Nem sempre o processo com implementação lexical é foneticamente abrupto, como já diziam Wang e Lien (1993: 348) e como parece ter sido o caso do alçamento. E nem sempre o processo com implementação lexical é tardio. Observamos que também nos estágios iniciais do processo de alçamento, assim como nas etapas tardias, a implementação é

lexical, conforme os estudos históricos aqui realizados. Em todas as etapas, existem itens que escapam à sistematização: não foram encontrados alçados quando teriam ambiente para tal (*covil*, por exemplo). Assim, se o alçamento pode ser descrito por uma regra “pós-lexical”, as regras “pós-lexicais” poderiam ter implementação lexical (o que inviabiliza o modelo).

Kiparsky (1995) critica a caracterização do número de traços fonéticos para a diferenciação das regras propostas por Labov (1994). Na reformulação de Kiparsky (1995), o processo de alçamento também não se enquadra. Argumenta o autor que a difusão lexical se comporta em todo aspecto como a analogia lexical. Esse seria um processo foneticamente abrupto, o que, mais uma vez, não se enquadra nas descrições históricas do fenômeno do alçamento, pois o processo de alçamento parece ter sido foneticamente gradual de acordo com Naro (1973: 14, já citado) e Teyssier (1997) e, portanto, não seria de se esperar que fosse implementado por difusão lexical, se difusão lexical fosse analogia lexical. E, ainda, segundo Phillips (1998), não é de se esperar que encontremos processos analógicos nos itens familiares ou nos mais freqüentes, mas nos menos freqüentes ou menos familiares, ou seja, quando a memória falha.

Não encontrei evidências de a implementação lexical estar presente nas regras lexicais apenas, mas também nas “pós-lexicais”. Assim, concordo com Phillips (1984) quando afirma que a implementação das mudanças é sempre lexical. Atuando quer nas etapas iniciais de um processo, quer nas finais, quer em processos foneticamente graduais, quer em processos foneticamente abruptos, embora muitas vezes a mudança atinja todo o léxico não deixando vestígio da atuação lexical.

Labov (1994: 604) nos diz que a distinção entre possíveis homônimos (*Sentido!* x *sintido*; *purção* x *porção*; *Peru* x *piru*) ocorre para se evitarem os mal-entendidos. Para ele, esta distinção não poderia ser consciente porque essa escolha demandaria tempo, e a escolha entre um item e outro tem que ser rápida no momento da comunicação. No caso do alçamento, a distinção entre os homônimos não é aleatória, como era de se esperar se fosse só para se evitarem os mal-entendidos. Os itens alçados são sempre os menos prestigiados

e os não alçados são sempre os mais prestigiados, como vimos. Esta é uma marcação estabelecida socialmente. Se a distinção entre os homônimos não é consciente no momento da comunicação, isto não implica em não haver ou ter havido uma marcação determinada socialmente nos itens lexicais. No caso do alçamento, por exemplo, a opção pelo alçamento recai, sistematicamente, nos itens populares, de uso menos prestigiado. Há uma marcação social, estabelecida historicamente, que estigmatiza o processo de alçamento, e também há uma marcação nos itens *usados* em situações consideradas menos prestigiadas. E a escolha não consciente não implica em não passar sentidos já marcados historicamente.

É óbvio que não podemos dizer que o falante teve a *intenção* de marcar a palavra com o alçamento nas etapas do alçamento “*from below*”, ou abaixo do nível de conscientização, mas a marcação estabelecida socialmente existiu (itens familiares, primeiro), marcando uma identidade social. Já nas etapas “*from above*”, ou acima do nível de conscientização, a intenção de marcação é evidente (*piscucinho*, *litrinha*, etc.). Se fizermos coincidir “produzir *sentido*” com “*intenção* consciente”, estamos, por definição, excluindo os processos “*from below*”, já que eles estão abaixo do nível de conscientização. E, então, não podemos averiguar se forma é sentido para esse tipo de processo. Não devemos, pois, fazer coincidir “produzir *sentido*” com “*intenção* consciente”. Existe um sentido estabelecido socialmente, que é passado de geração a geração, e que não é necessariamente consciente.

Além disso, devemos ter bastante critério quando utilizarmos o “princípio uniformitário”, ou seja, “o uso do presente para explicar o passado”, pois as marcações do presente podem não ser exatamente as mesmas marcações estabelecidas em todo o processo. No passado, podemos ter julgamentos, valorações sociais diferentes das marcações do presente. O princípio uniformitário se aplica à estrutura gramatical.

O alçamento se iniciou nos itens adquiridos, populares, cotidianos, menos marcados, posteriormente adquiriu certo grau de estigma social, e o processo passou a atuar nos itens mais pejorativos, palavrões, itens jocosos, irônicos.

É interessante observar aqui que recortes diferentes na amostragem dos dados podem nos levar a explicações diferentes do processo. Em Viegas (1987) temos uma amostragem baseada em dados colhidos em entrevistas gravadas em determinadas comunidades de fala e podemos observar alguma regularidade que subjaz ao processo, acrescida de exceções. Esta regularidade pode ficar comprometida ou tendenciada se o número de ocorrências de determinado item for muito grande.

A regra de harmonização proposta para o /e/ em Viegas (1987) parece se encaixar bem nos estudos históricos como uma tendência, ou, nos termos de Bybee (2001), um esquema que emerge do léxico. Já a regra proposta em Viegas (1987) para o /o/ – levantamento devido às consoantes adjacentes – é muito abrangente e dificilmente encontramos itens que aí não se enquadram, ou seja, atinge quase todo o léxico (como ocorreu em Portugal), esquema que aqui parece ter sido “barrado”.

Abaurre-Gnerre (1981) não dá ênfase, no caso de /e/, ao processo de harmonização vocálica, embora o processo de harmonização encontre confirmação nos estudos históricos. Na sua amostragem, a autora incluiu itens como pequeno e melhor (analisados separadamente aqui), conduzindo sua análise, tanto para o /o/, quanto para o /e/, para o levantamento devido às consoantes adjacentes.

Se analisássemos itens lexicais isolados, poderíamos dizer que a questão é puramente lexical, pois, se compararmos itens com ambientes semelhantes, um alça e o outro não. Oliveira (1991) inclui nos seus dados empréstimos que podem ter vindo com a vogal já alta, e não explica por que muitos itens comuns em estilo informal e sem ambiente favorecedor alçam e outros não.

Estamos vendo que a questão pode ser assim interpretada: existem ambientes favorecedores e restrições lexicais relacionadas ao grau de familiaridade do item e à valoração social do item em uso.

Oliveira (1992) faz a postulação do contexto fonético atuando “a posteriori”, como um fixador das mudanças. Segundo o autor, as mudanças que “vingam” seriam estabelecidas por relações harmônicas, que seriam processos como a assimilação. Relações desarmônicas,

como a dissimilação, não estariam previstas na sua formulação de um modelo de mudança.

A grande dificuldade foi encontrar um modelo que compatibilizasse a influência dos ambientes favorecedores e a implementação lexical para todo tipo de mudança. O modelo de léxico com armazenagem exemplar, proposto em Bybee (2001) e em texto inédito, parece enquadrar bem o processo de alçamento:

A literatura a respeito do efeito da frequência na difusão lexical mostra que até mesmo as mudanças foneticamente graduais, que são destinadas a serem lexicalmente regulares em muitos casos, mostram difusão lexical quando estão em progresso.

“Eu argumento que mudanças foneticamente condicionadas que afetam palavras com alta frequência antes das palavras com frequência baixa são melhor enquadradas em um modelo exemplar de representação fonológica que permite à mudança ser foneticamente e lexicalmente gradual.”

“Uma propriedade importante do modelo exemplar é a ênfase nas palavras como unidades de armazenamento.”

“... a mudança sonora ocorre em tempo real quando as palavras são usadas e seus efeitos são registrados na memória, produzindo uma mudança gradual nos itens lexicais baseada na experiência real do falante com os itens.” (BYBEE (inédito))⁹

Não observei a frequência dos itens, mas atribuí à familiaridade dos itens o poder de mudança. Fidelholtz (1975) atribui à alta frequência dos itens o favorecimento de mudanças do tipo do alçamento, mas explica que certos itens não muito frequentes podem ser atingidos pelo processo caso sejam itens familiares. Podemos observar hoje o favorecimento dos palavrões e xingamentos, talvez não tão frequentes, em processos como o de alçamento e como o de cancelamento do /d/ em gerúndios, por exemplo (Viegas, inédito). Existe, pois, uma questão de valoração social atuando na marcação dos itens, assim como a questão frequencial do item relacionada ao grupo social que está “levando” a mudança.

Desse modo, poderíamos até manter o fator regularidade dos neogramáticos como sendo consistente com um processo do tipo do

alçamento (pelo menos no caso do e), mas não o fato de a mudança atingir o léxico abruptamente, sem exceções. A mudança é lexical e não está condicionada por fatores fonéticos apenas.

Poderíamos perguntar: forma é sentido? Se entendermos sentido como “valor apreciativo” como em Bakhtin (1986), então, a resposta é afirmativa. Assim, forma é função. Função de marcar a identidade de um grupo social nas mudanças que se iniciam “*from below*” (nos termos de LABOV, 1994), principalmente nos grupos não prestigiados, e função de estabelecer marcação de prestígio ou de estigma nas mudanças mais marcadas que se iniciam “*from above*”.

A noção de que há uma “decadência lingüística”, vulgarmente associada à mudança lingüística, pode decorrer também da valoração social dada às mudanças que se iniciam “*from below*”, atingindo os itens familiares primeiro – os itens adquiridos (em oposição aos aprendidos).

Concluo, utilizando a terminologia de Labov (1994) para, de modo geral, esquematizar a implementação dos processos de mudança sonora, salientando que existe, na realidade, uma gradação entre processos “*from below*” e “*from above*”:

I - Processos que se iniciam “*from below*”:

ETAPAS INICIAIS: “intimidade”, familiaridade com os itens ou “não-marcas evidentes” nos itens familiares: *permitem* a atuação do processo até esse atingir os mais marcados.

ETAPAS POSTERIORES: a) marcas desprestigosas: *motivam* a atuação do processo; (se o processo tornou-se b) marcas de prestígio: *inibem* a atuação do processo. estigmatizado (“*from above*”))

II – Processos prestigiados que se iniciam “*from above*”:

ETAPAS INICIAIS: a) marcas de prestígio: *motivam* a atuação do processo; b) marcas de desprestígio: *inibem* a atuação do processo.

ETAPAS POSTERIORES: “não-marcas evidentes” nos itens: *permitem* a atuação do processo, podendo até atingir os itens menos prestigiados.

REFERÊNCIAS BIBLIOGRÁFICAS

ABAURRE-GNERRE, M. B. Processos fonológicos segmentais como índices de padrões diversos nos estilos formal e casual do Português do Brasil. *Cadernos de Estudos Linguísticos*, Campinas: Universidade Estadual de Campinas, n. 2, p. 23-45, 1981.

BAKHTIN, M. (VOLOCHINOV). *Marxismo e filosofia da linguagem*. 3. ed. Trad. Michel Lahud e Yara Frateschi Vieira. São Paulo: HUCITEC, 1986.

BLUTEAU, D. R. de. *Vocabulario portuguez e latino*. Coimbra: Collegio das Artes da Cia de Jesu, (1712-1721). (8 volumes, 2 tomos cada).

BYBEE, J. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change, to appear in *Language Variation and Change*.

BYBEE, J. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001

BYBEE, J. Regular Morphology and the Lexicon. *Language and cognitive processes*, U.K., Erlbaum, 10 (5), p. 425-455, 1995.

CARTAS do Brasil e mais Escritos do P. Manuel da Nóbrega – Introdução Novas Histórias e Críticas: Serafim Leite S. I. Belo Horizonte: Itatiaia, 2000. Edição fac-similar comemorativa dos 500 anos da Descoberta do Brasil (Coimbra por ordem da Universidade, 1955).

COROMINAS, J. *Dicionário Crítico etimológico de la lengua castellana*. Madrid, 1954-1957. 4 vols.

CUNHA, A. G. da. *Dicionário etimológico nova fronteira da língua portuguesa*. Rio de Janeiro: Nova Fronteira, 1982.

FERREIRA, A. B. de H. *Novo Dicionário da Língua Portuguesa*. Rio de Janeiro: Nova Fronteira, 1975.

FIDELHOLTZ, J. Word frequency and vowel reduction in English. *CLS* 11, p.200-213, 1975.

KATO, M. A. Aquisição e aprendizagem da língua materna: de um saber inconsciente para um saber metalingüístico. In: CABRAL, L.; MORAIS, J. (Org.). *Investigando a linguagem*. Ensaios em homenagem a Leonor Scliar Cabral. Florianópolis: Mulheres, 1999.

KIPARSKY, P. The Phonological Basis of Sound Change. In: *The Handbook of phonological theory*. Cambridge, Mass.: MIT Press, 1995.

LABOV, W. *Principles of Linguistic Change: social factors*. Oxford: Black Well, 2001. Vol. 2.

- LABOV, W. *Principles of Linguistic Change: internal factors*. Oxford: Black Well, 1994. Vol. 1.
- MAIA, C. História do galego-português – Estudo lingüístico da Galiza e do Noroeste de Portugal desde o século XIII ao século XVI (com referência à situação do galego moderno). Coimbra: I.N.J.C., Fundação Calouste Gulbenkian, 1986.
- MATTOS E SILVA, R. *O português arcaico: fonologia*. São Paulo/Bahia: Contexto/UFBA, 1991.
- MILROY, L. *Language and Social Networks*. Oxford: Black Well, 1987.
- NARO, A. H. *Estudos diacrônicos*. Petrópolis: Vozes, 1973.
- OLIVEIRA, M. A. de. O léxico como controlador das mudanças sonoras. *Revista de Estudos da Linguagem*, FALE/UFMG, Belo Horizonte, 1995.
- OLIVEIRA, M. A. de. Aspectos da difusão lexical. *Revista de Estudos da Linguagem*, FALE/UFMG, Belo Horizonte, p. 31-41, 1992.
- OLIVEIRA, M. A. de. The neogrammarian controversy revisited. *International Journal of the Sociology of Language*, Berlin, v. 89, 1991.
- PAUL, H. *Princípios fundamentais da história da língua*. Lisboa: Fundação Calouste Gulbenkian, 1966.
- PHILLIPS, B. S. Lexical Diffusion is *not* lexical analogy. *Word*, 49 (3), p. 369-381, 1998.
- PHILLIPS, B. S. Word frequency and the actuation of sound change. *Language* 60 (2), p.320-342, 1984.
- TAVARES, J. P. *Antologia de textos medievais - selecção, introdução e notas*. 2. ed. Lisboa: Livraria Sá da Costa ed., 1961.
- TEYSSIER, P. *História da língua portuguesa*. Trad. Celso Cunha. São Paulo: Martins Fontes, 1997.
- VIEGAS, Maria do Carmo. Evidências em favor de um léxico com armazenamento exemplar. (Inédito)
- VIEGAS, M. do C.. O alçamento de vogais médias pretônicas e as conseqüências de diferentes recortes na amostragem. In: *Anais do II Seminário Internacional de Fonologia*. Porto Alegre: PUC-RS. (No prelo)
- VIEGAS, M. do C. *O alçamento das vogais médias pretônicas e os itens lexicais*. 2001. Tese (Doutorado em Estudos Lingüísticos) – Faculdade de Letras da UFMG, Belo Horizonte.

VIEGAS, M. do C. *Alçamento de vogais médias pretônicas: uma abordagem sociolingüística*. 1987. Dissertação (Mestrado em Estudos Lingüísticos) – Faculdade de Letras da UFMG, Belo Horizonte.

WANG, W.S.-Y; LIEN, C. Bidirectional diffusion in sound change. In: *Historical Linguistics: problems and perspectives*. Londres: Longman, 1993.

WANG, W.S.-Y.; CHENG, C.C. Implementation of Phonological Change: the Shuang-feng Chinese Case. In: WANG, W. S-Y (Ed.). *The lexicon in phonological change*. The Hague: Mouton, 1977.

WILLIAMS, E. B. *Do latim ao português*, fonologia e morfologia históricas da Língua Portuguesa. Trad. Antônio Houaiss. 3. ed. Rio de Janeiro: Tempo Brasileiro, 1975.

A importância da proeminência da frase fonológica no português brasileiro*

Luciani TENANI
UNESP/SJRP

Abstract

In order to show the importance of phonological phrase prominence in Brazilian Portuguese, we make an intonational analysis of contexts which characterize vowel merge and vowel deletion processes. We will argue that this phonological prominence blocks vowel sandhi processes because it carries the syntactic notion of recursivity in Portuguese. The results contribute to syntax-phonology interface discussion.

1. INTRODUÇÃO

Neste trabalho, são analisados processos de sândi externo em que uma das vogais da seqüência vocálica é acentuada, seja essa seqüência de vogais iguais ou diferentes. Uma descrição dos contextos relevantes para nossa discussão é dada na próxima seção, ao retomarmos as análises feitas para o sândi vocálico no português brasileiro (doravante PB). Nosso objetivo é encontrar evidências entoacionais da importância da preservação da proeminência da frase fonológica (doravante ϕ) em PB, quando há um conflito entre a aplicação de processo segmental e a preservação de proeminência prosódica. Para alcançar esse objetivo, é adotada a abordagem da fonologia prosódica e da fonologia entoacional. Na terceira seção deste texto, apresentamos sucintamente essas abordagens e o *corpus* analisado que foi construído com base na proposta da chamada fonologia laboratorial. Na quarta seção, discutimos os resultados obtidos com vistas a tratar das pistas que sinalizam como se dá a interpretação fonológica da direção da recursividade sintática no PB. Na última seção, a partir dos resultados alcançados, tecemos considerações acerca da interface sintaxe-fonologia.

2. O SÂNDI VOCÁLICO E A IMPORTÂNCIA DA FRASE FONOLÓGICA

O sândi vocálico externo se caracteriza pelo encontro de duas vogais que pertencem a vocábulos diferentes. Se as vogais são iguais, configura-se a degeminação; se as vogais são diferentes, configura-se a elisão ou a ditongação. Um panorama completo das possibilidades de aplicação e bloqueio da degeminação (doravante DG), da elisão (doravante EL) e da ditongação (doravante DT) é traçado por Bisol (1993) que toma por base os dados do projeto NURC. Resumidamente,

pode-se afirmar que o acento é o fator que gera os contextos de bloqueio do sândi vocálico, pois: (i) quando ambas as vogais são tônicas, nenhum dos três processos se implementam, como exemplifica (1.1); (ii) quando a primeira vogal é tônica, como em (1.2), a elisão não ocorre, como exemplifica (1.2.ii); (iii) quando a segunda vogal é tônica, como em (1.3), nem a elisão nem a degeminação ocorrem, como exemplificam (1.3.i) e (1.3.ii).¹

- | | | |
|-----|-----------------------|---------------|
| (1) | 1.i. Vi isso | [vi isu] |
| | 1.ii. Dá isso | [da isu] |
| | 1.iii. Vi este | [vi estʃɪ] |
| | 1.2.i. Babá amorosa | [babamoroza] |
| | 1.2.ii. Babá elegante | [babaelegãʃɪ] |
| | 1.2.iii. Bebê urinado | [bebeurinadu] |
| | 1.3.i. Como uvas | [komuuvas] |
| | 1.3.ii. Coma uvas | [komauvas] |
| | 1.3.iii. Come uvas | [komnuvas] |

A partir da descrição de Bisol (1993), consideramos, neste texto, os contextos de bloqueio de sândi que são motivados pela presença do acento e também controlamos a fronteira prosódica tomando por base a argumentação de Abaurre (1996). Essa pesquisadora retoma do trabalho de Bisol (1993) os contextos em que a segunda vogal é acentuada tanto para o contexto segmental da EL, em (2.1), quanto da DG, em (2.2), e observa que, em ambos os casos, o ambiente em questão é o mesmo: *a segunda vogal da seqüência é portadora de acento nuclear no nível da frase fonológica*. No entanto, em (2.1.ii) e (2.2.ii), os processos não se aplicam.

- | | |
|-----|--|
| (2) | 1.i. [[Ele] φ [compra] φ [uvas caras] φ] I ⇒ Ele cómpr[u]vas caras |
| | 1.ii. [[Ele] φ [compra] φ [uvas] φ] I ⇒ *Ele compr[ú]vas |
| | 2.i. [[Eu] φ [compro] φ [uvas sempre] φ] I ⇒ Eu cómpr[u]vas sempre |
| | 2.ii. [[Eu] φ [compro] φ [uvas] φ] I ⇒ *Eu compr[ú]vas. |

Para Abaurre (1996: 45), tais fatos indicam que a análise feita por Bisol (1993),

embora descritivamente adequada e perfeitamente coerente com a direção *bottom-up* de atribuição de acento e construção da grade métrica por ela assumida, não parece realmente explicar por que a elisão e a degeminação se tornam possíveis quando o acento primário de palavra é, em suas palavras, “enfraquecido por razões prosódicas e rítmicas”.

Ao levar em conta dados como os apresentados em (2), a autora afirma que

o bloqueio a esses processos ocorre quando [o] acento primário de palavra, atribuído no componente lexical, é também interpretado, pós-lexicalmente, *como acento frasal, portador de informação sintática*, dentro de uma hierarquia de proeminências prosódicas sintaticamente motivadas. (ABAURRE, 1996: 46)²

Esse fato está relacionado, segundo a pesquisadora, ao algoritmo de formação de ϕ^3 na medida em que a proeminência relativa desse domínio sinaliza a direção da recursividade sintática nas línguas naturais (o que é, por definição, um parâmetro sintático). Como o Português é uma língua de recursividade à direita (cf. CINQUE, 1993), ao acento mais à direita será atribuído o valor *forte*, assim, o acento de ϕ mais à direita tem relação com a recursividade sintática.

Recorrendo ao trabalho de Nespor (1994), Abaurre (1993) traz informações sobre experimentos em aquisição da linguagem que parecem confirmar a saliência prosódico-sintática do acento de frase fonológica e sua relevância no estabelecimento da interface entre os módulos fonológico e sintático. Também Hayes (1991) afirma que a atribuição de acento no nível frasal exhibe diferenças marcantes com relação à atribuição do acento de palavra, as quais parecem corroborar uma interpretação do acento frasal como tendo uma natureza não puramente fonológica. Os resultados desses trabalhos são tomados como a base para a hipótese formulada por Abaurre (1996), segundo a qual “a implementação dos processos fonológicos, como o sândi

externo no português, pode ser (em parte) regulada pelo acento frasal” (p. 47).

Abaurre (1996) ainda fundamenta o fato de o acento frasal estar relacionado ao padrão de proeminências rítmicas nas afirmações feitas por Selkirk (1984). A pesquisadora brasileira vincula explicitamente os possíveis padrões de proeminências rítmicas dos enunciados às suas propriedades entoacionais, quando cita Selkirk (1984):

os elementos tonais acidentalmente relevantes (*pitch accents*) são atribuídos às palavras na estrutura superficial independentemente dos padrões de acento frasal, e os padrões frasais são parcialmente definidos como uma função da localização das palavras que, na sentença, são portadoras do *pitch accent*. (ABAURRE, 1996: 47).

Em outras palavras, os padrões rítmicos são determinados pelos padrões entoacionais, e esta relação pode definir, segundo Selkirk (1984), o lugar da proeminência relativa de constituintes sintáticos, através de sua delimitação prosódica.

Dessas considerações, torna-se evidente a importância do acento frasal e do seu estatuto prosódico, pois a este acento estão associadas informações não apenas de natureza rítmica, mas também entoacional e, por fim, sintáticas. Como o acento frasal é aquele que bloqueia os processos de sândi em PB, Abaurre (1996) interpreta esse fato como um conflito de natureza ‘teleológica’ ao afirmar que:

Esse conflito envolve, por um lado, princípios de boa-formação das seqüências fonológicas que se manifestam através da *tendência à otimização da cadeia silábica* [...]; envolve ainda, por outro lado, *a tendência a preservar a estrutura portadora da informação sintática relevante*. (p. 48)

Nesse embate, a análise dos dados da DG e da EL indicia que a tendência para se preservar a estrutura portadora de informação sintática prevalece sobre a tendência à otimização silábica no nível da frase fonológica. Nessa afirmação de Abaurre (1996) está implícita outra a respeito da interface dos módulos da gramática: a sintaxe tem prevalência sobre a fonologia.

Abaurre (1996) conclui seu texto observando que Bisol (1996a) não deixa de perceber que o acento de frase fonológica está envolvido no bloqueio da DG e da EL, porém deixa de formalizar o fato considerado relevante que é a referência à proeminência de nível hierarquicamente superior ao acento lexical, ou seja, o acento da frase fonológica. Dessa maneira, Abaurre (1996) argumenta que é a frase fonológica o domínio prosódico de aplicação e bloqueio dos fenômenos de sândi e, desse modo, o acento deste domínio deve ser mantido, o que mostra que, em PB, a necessidade de se preservarem ao máximo os contextos estruturais portadores de informação sintática prevalece sobre a otimização no nível silábico.

Posto que a DG e a EL são bloqueadas pelo elemento mais proeminente de ϕ , tem-se aí uma evidência da relevância desse domínio para o PB. Avançamos, porém, neste trabalho ao colocar em cena o choque de acentos. Por meio de experimento, a ser apresentado na próxima seção, controlamos a distância entre os acentos das palavras sujeitas ao sândi e a relação de proeminência entre os acentos no domínio de ϕ . O controle dessas variáveis se justifica em função dos fatos relativos ao choque de acentos no Português.

A literatura que trata do choque de acentos, especialmente os trabalhos que se inserem no quadro teórico da fonologia métrica, aponta a distância entre os acentos e a estrutura prosódica como dois fatores que estão relacionados às restrições rítmicas identificadas nas línguas em geral.⁴

As configurações das grades métricas geradas desse mapeamento nem sempre são bem-formadas. Em alguns casos, as proeminências estão tão próximas que resultam em um choque de acentos (*stress clash*); em outros casos, as proeminências estão tão longe umas das outras que resultam em ‘lapsos’ de acentos (*stress lapses*). Neste trabalho, a atenção é dada apenas a casos em que se configuram choque de acentos. Para solucionar essas configurações de modo a implementar uma seqüência com uma maior alternância rítmica, regras de reajustes são aplicadas. As estratégias de resolução de choque de acentos e o nível em que esse choque é avaliado como uma configuração malformada

diferem entre as línguas. Além dessas duas variáveis, a direção (direita/esquerda) em que operam os ajustes rítmicos também varia entre as línguas.

O trabalho sobre choque de acentos de Abousalh (1997) mostra que, em PB, a resolução da colisão acentual opera dentro do domínio de ϕ por meio de uma estratégia de retração à esquerda do acento. Já o choque de acentos entre ϕ_s parece não ser resolvido por essa mesma estratégia, segundo a pesquisadora. Esse fato pode ser indício de que, provavelmente, a língua deve dispor de algum outro mecanismo para resolver essa configuração rítmica. A autora aborda como estratégia de resolução dos choques apenas 'o apagamento/movimento de acento, a inserção de pausas e de contornos entoacionais' (p. 19), pois opta por trabalhar 'somente com uma transcrição orientada por critérios lingüísticos' (p. 18), isto é, a análise é baseada em uma transcrição auditiva de dados não obtidos em laboratório. Portanto, não são investigadas outras estratégias como a inserção de uma batida rítmica (*beat insertion*) entre os acentos ou os correlatos acústicos das estratégias empregadas em PB.

Também Sandalo & Trunckenbrodt (2002) tratam do choque de acentos em PB e tomam por base julgamentos dos falantes sobre a possibilidade de retração de acentos em sentenças lidas (como enunciados de jornal, por exemplo), mas também não é aventada a possibilidade de haver outras estratégias de resolução de choque acentual.⁵

Por ora, pode-se afirmar que não há dúvidas da relevância de ϕ quer para o bloqueio da DG e da EL, quer para o bloqueio da resolução de choque de acentos. Neste texto, buscamos verificar a interação entre o bloqueio desses processos de sândi e as estratégias de resolução de choque de acentos.⁶ Para alcançar esse objetivo, são consideradas, na próxima seção, as variações da F0 analisadas em termos de eventos tonais associados aos elementos proeminentes de ϕ e da frase entoacional (doravante *I*).

3. ABORDAGEM TEÓRICA

O algoritmo de formação da frase fonológica que adotamos é aquele inicialmente formulado por Nespor & Vogel (1986) e que foi apresentado por Abaurre (1996:46) do modo que é dado em (3). O algoritmo em (3) sofreu inúmeras críticas (cf. por exemplo, INKELAS & ZEC, 1990; 1995) e, em razão dos problemas identificados, são feitas várias propostas alternativas, inclusive para o português (cf. por exemplo, GUIMARÃES (1998); SANDALO & TRUNCKENBRODT (2002)). Optamos por não discutir, neste trabalho, esse aspecto relativo à formação desse constituinte, embora esse seja um ponto extremamente importante para a discussão sobre a interface sintaxe-fonologia. Para a frase entoacional, também adotamos o algoritmo proposto por Nespor & Vogel (1986: 189; 191), expresso em (4).

(3) Formação da frase fonológica ϕ

i. Domínios de ϕ

O domínio ϕ consiste de um C (grupo clítico) que contém uma cabeça lexical (X) e todos os C's no seu lado não-recursivo até o C que contém outra cabeça fora da projeção máxima de X.

ii. Construção de ϕ

Reúna em uma ϕ n-ária todos os C's incluídos em uma cadeia delimitada pela definição do domínio ϕ .

iii. Proeminência relativa de ϕ

Em línguas cujas árvores sintáticas ramificam-se à direita, o nódulo de ϕ mais à direita é rotulado como *forte*; em línguas cujas árvores sintáticas ramificam à esquerda, o nódulo de ϕ mais à esquerda é rotulado como *forte*. Todos os nódulos irmãos de *forte* são rotulados de *fracos*.

(4) Formação da frase entoacional I

i. Domínios de I

Um domínio I pode consistir de: (a) todos os ϕ_s em uma cadeia que não é estruturalmente relacionada à árvore da sentença no nível da estrutura S; ou (b) qualquer seqüência remanescente de ϕ_s adjacentes a uma sentença raiz.

ii. Construção de *I*

Reúna em um *I* de ramificação n-ária todos os ϕ_s incluídos em uma cadeia delimitada pela definição do domínio *I*.

iii. Proeminência de *I*

Dentro de *I*, um nóculo é rotulado como *forte* com base na sua proeminência pragmática; todos os outros nóculos são rotulados como *fraco*.

A análise da variação de F0 em termos de eventos tonais é feita com base na abordagem da fonologia entoacional, que segue a proposta de Ladd (1996) e Frota (1998). Essa abordagem se insere na linha de análise da entoação em níveis tonais inaugurada por Pierrehumbert (1980).

A perspectiva de análise da entoação que assumimos é a denominada por Frota (1998: 11) como 'visão integrada' que tem por característica considerar as propriedades entoacionais como uma das pistas da estrutura prosódica. Nessa abordagem, apenas a altura (*pitch*) é tomada como traço entoacional característico. O acento (*stress*) e a junctura (*junction*), dois outros aspectos relevantes dentro da tradição de análise entoacional a que se vincula essa abordagem, decorrem (de algum modo) da estrutura prosódica. Vistos dessa maneira os aspectos básicos que envolvem os contornos entoacionais, a abordagem da entoação adotada dá à estrutura prosódica um papel fundamental.

Desse modo, essa visão integrada da entoação se opõe àquelas em que é dado à entoação um *status* independente da estrutura prosódica, como o faz Selkirk (1984),⁷ e àquelas em que a constituição dos domínios é dada pela entoação, como o fazem Beckman & Pierrehumbert (1986), Pierrehumbert & Beckman (1988).⁸

Essa visão da entoação é formalizada pela teoria autosegmental e métrica da fonologia entoacional (cf. cap. 2 de LADD, 1996) que propõe que a entoação tenha uma organização fonológica própria. Uma das asserções fundamentais dessa teoria é que a frequência fundamental (F0) deve ser interpretada como uma seqüência de eventos fonológicos discretos, e não como um contorno contínuo variável que pode ser caracterizado segundo sua forma ou direção

(cf. LADD, 1992: 322). Dois são os tipos de eventos – denominados de tonais por terem como unidade básica o tom – suficientes para descrever as variações de F0: os acentos tonais (*pitch accents*) e os tons de fronteira (*boundary tones*) – cf. Hayes & Lahiri (1991: 50).

O primeiro tipo de evento tonal é definido como o tom, que é associado à sílaba acentuada, cuja proeminência é (de algum modo) definida independentemente do contorno entoacional. Formalmente, eles são indicados por um asterisco (ex. H*, L*). Esses tons podem ser simples (alto: H*, ou baixo: L*), quando formados por um tom, ou complexos, quando compostos por dois tons (LH*, L*H, HL*, H*L). O evento bitonal pode ser entendido nos seguintes termos: *“the starred tone falls on a stressed syllable, while the other tone specifies a rapid pitch change just before or after the stress”* (HAYES & LAHIRI, 1991: 51). O segundo tipo de evento tonal, os tons de fronteira, é formalmente indicado por L% ou H% (cf. HAYES & LAHIRI, 1991, LADD, 1996, entre vários outros) ou por Li ou Hi (cf. FROTA, 1998, 2001). Como o próprio nome sugere, esses tons são ligados a fronteiras de constituintes prosódicos e não a sílabas propriamente ditas, ou seja, a mudança de altura tem como alvo uma fronteira de constituinte (cf. HAYES & LAHIRI, 1991: 51).

A partir da visão integrada entre entoação e constituintes prosódicos, empreendemos, na quarta seção, uma descrição precisa de como os eventos tonais estão associados ao domínio ϕ e como essa configuração traz pistas de como a fonologia interpreta informação sintática. Na próxima seção, detalhamos o experimento realizado explicitando as variáveis controladas.

3.1. O experimento

Com o objetivo de obter evidências da relevância do domínio ϕ em PB, elaboramos contextos em que há interação entre o bloqueio da DG e da EL e o choque de acentos das vogais candidatas ao sândi. A fim de alcançar esse objetivo, foi feito um experimento em que são controlados os seguintes fatores: a qualidade das vogais, a tonicidade das vogais, a distância entre os acentos e as fronteiras prosódicas.

Cabe observar que a tonicidade das vogais sujeitas ao sândi é controlada de modo a relacionar essa tonicidade à saliência prosódica de ϕ , domínio relevante para o bloqueio do sândi vocálico e para a resolução de choque acentual. Também consideramos a fronteira de *I*, pois nos interessa verificar se essa fronteira exerce algum papel em contextos de sândi em PB.⁹ Ou seja, foram elaborados contextos em que o acento de palavra é também interpretado como acento frasal, o qual se manifesta de modo mais evidente por meio da variação de F0 (cf. MORAES, 1987; MASSINI-CAGLIARI, 1992).¹⁰ Desse modo, foram elaboradas sentenças nas quais a localização das fronteiras de ϕ e de *I* e os contextos de sândi foram sistematicamente variados.

Abaixo, são dados alguns exemplos de sentenças do *corpus*, com indicação dos contextos de sândi (sublinhado), dos acentos (em caixa alta) e das fronteiras prosódicas relevantes.

- (5) 1. [a ALUNA ÁRABE] ϕ [enviou uma carta] ϕ [à cantora] ϕ
 2. [a ALUNA] ϕ [AGE] ϕ [com discrição] ϕ [em público] ϕ
 3. [a ALUNA] ϕ [AGE SEMPRE] ϕ [com discrição] ϕ [em público] ϕ
 4. [a ALUNA,] *I* [Ávida por justiça,] *I* [falou com a diretora.] *I*

Em todas as sentenças em (5), o artigo *a* pertence ao mesmo ϕ do nome *aluna* e, por essa razão, espera-se a aplicação da DG, uma vez que ϕ é o domínio preferencial para aplicação de sândi externo (porque sempre ocorre a aplicação desse processo nesse domínio entre vogais átonas – cf. resultados de BISOL, 1993 e TENANI, 2002). Em (5.1), *aluna árabe* é formada por ϕ_s passíveis de sofrerem reestruturação, conforme previsto pelo algoritmo desse domínio, constituindo-se, assim, em uma unidade fonológica pelo fato de *árabe* constituir um ϕ não-ramificado que é complemento da cabeça lexical X (no caso, *aluna*) do seu lado recursivo. Em (5.2), a relação é entre as fronteiras de ϕ , sendo os dois ϕ_s não-ramificados. Em (5.3), a relação prosódica relevante se dá entre um ϕ não-ramificado, formado por um nome em posição de sujeito (*aluna*), e outro ϕ -ramificado, formado por um verbo e um advérbio (*age sempre*). Cabe observar que, pelo algoritmo de ϕ , *age sempre* é mapeado em dois ϕ_s [*age*] ϕ [*sempre*] ϕ que são passíveis de reestruturação [*age sempre*] ϕ .

Estamos considerando que a estrutura relevante é [*age sempre*]φ. Finalmente em (5.4), a fronteira prosódica considerada é do domínio *I*. Nesse caso, é prevista a fronteira de *I* entre *aluna* e *ávida* em razão de haver o mapeamento do parêntese ‘*ávida por justiça*’ como um único *I*, o que a delimita em relação à porção anterior e posterior da sentença em que é encaixada. Todas as fronteiras prosódicas consideradas são elencadas abaixo.

(6) *Tipo de relação prosódica*

- a. Mesmo φ
- b. φ+φ-não-ramificado
- c. φ+φ-ramificado
- d. I+I

Como já dissemos, foi controlada a tonicidade das vogais sujeitas ao sândi de modo a relacionar essa tonicidade à saliência prosódica do domínio da frase fonológica. Com base nos trabalhos de Bisol (1993, 1996a, b, 2000) e de Abaurre (1996), constata-se que a tonicidade da segunda vogal é crucial, pois, é o acento dessa vogal que gera o bloqueio da DG e da EL quando esse for interpretado como acento de φ. Dessa maneira, foram considerados contextos em que a segunda vogal é tônica (v+v’) tanto para a seqüência de /a+á/, como em (7.i), que permite observar o processo da DG, quanto para seqüência de /a+ú/, como em (7.ii), que caracteriza o contexto segmental da EL em PB.

- (7) 1.i. [a ALUNA Árabe] φ [enviou] φ [uma carta] φ [à cantora] φ
 1.ii. [a ALUNA Útil] φ [ganhou] φ [uma viagem]φ [de férias]φ
 2.i. [a ALUNA] φ [Age] φ [com discrição] φ [em público] φ
 2.ii. [a ALUNA] φ [usa] φ [chapéu branco]φ [em público]φ
 3.i. [a ALUNA] φ [Age SEMpre] φ [com discrição] φ [em público] φ
 3.ii. [a ALUNA] φ [usa SEMpre] φ [chapéu branco]φ [em público]φ
 4.i. [a ALUNA]I[Ávida por justiça]I [falou com a diretora]I
 4.ii. [a ALUNA]I[Útil mas desorganizada]I [encantou o público brasileiro]I

Optamos por incluir em nossa investigação o contexto em que o acento recai na primeira sílaba da seqüência da DG ou da EL e consideramos as seqüências /á+a/ (cf. 8.i) e /á+u/ (cf. 8.ii) nas diferentes estruturas prosódicas. A consideração desses contextos completa o quadro descritivo e fornece elementos que corroboram a visão de que é o acento mais à direita do domínio prosódico relevante que bloqueia o processo segmental.

- (8) 1.i. [o marajÁ aFOITO] φ [enviou] φ [uma carta] φ [à cantora] φ
 1.ii. [o marajÁ HUMILde] φ [recebeu flores] φ
 2.i. [o marajÁ] φ [ACEITA] φ [o papel] φ [de bandido]φ
 2.ii. [o marajÁ] φ [USAVA] φ [roupa branca] φ
 3.i. [o marajÁ] φ [APENAS ENVIou] φ [flores] φ [à bailarina]φ
 3.ii. [o marajÁ] φ [USAVA SEMPRE] φ [roupa clara]φ
 4.i. [o marajÁ]I [APÓS o CONFLITO]I [repousou em seu quarto]I
 4.ii. [o marajÁ]I [HUMILde como a cantora]I [recebeu os presentes]I

Ainda relacionada à tonicidade das vogais sujeitas ao sândi, foi sistematicamente controlada a distância entre os acentos das palavras sujeitas ao processo, com o objetivo de verificar se o processo é bloqueado apenas quando sua aplicação gera choque de acentos, como em (9.i). Essa distância foi considerada em termos de número de sílabas: uma ou duas sílabas entre os acentos primários. Ao ser controlada a distância entre os acentos, busca-se avaliar se há relação entre bloqueio/aplicação de sândi e a configuração métrica obtida. A partir das variáveis relacionadas ao acento das vogais foram elaboradas seqüências em que: (i) a distância entre os acentos é ora de uma, ora de duas sílabas átonas (respectivamente, 9.i *versus* 9.ii); (ii) ora a primeira vogal candidata ao sândi é átona e a segunda tônica (cf. 9.1 e 9.3), ora a primeira vogal candidata ao sândi é tônica e a segunda átona (cf. 9.2 e 9.4).

- (9) 1.i. [a ALUNA ÁRabe]φ
 1.ii. [a ASTRÓloga ÁRabe]φ
 2.i. [o marajÁ aFOITO]φ

- 2.ii. [o marajÁ africano]φ
- 3.i. [a aluna]φ [Age]φ
- 3.ii. [a astrÓloga]φ [Age]φ
- 4.i. [o marajÁ]φ [aceita]φ
- 4.ii. [o marajÁ]φ [aceitou]φ

Em razão dos objetivos colocados, também foi necessário levar em conta a variável distância entre os acentos de palavra em seqüências de vogais iguais em que ambas são átonas /a+a/. No contexto segmental considerado, a distância entre os acentos foi de duas, três e quatro sílabas, como exemplificam respectivamente (10.i), (10.ii) e (10.iii). Mais uma vez, foram controladas as fronteiras prosódicas (cf. 6) entre as vogais sujeitas ao sândi, como exemplificam (10.1) a (10.4). A consideração desse contexto em que as vogais são átonas visa a observar se há relação entre a configuração dos eventos tonais e a distância entre os acentos em diferentes estruturas prosódicas. Desse modo, buscamos comparar esses contextos em que as vogais são átonas àqueles em que pelo menos uma das vogais é acentuada.

- (10) 1.i. [a aluna afoita] φ [enviou] φ [uma carta] φ [à cantora] φ
 1.ii. [a aluna africana] φ [enviou] φ [uma carta] φ [à cantora] φ
 1.iii. [a astrÓloga africana]φ [enviou] φ [uma carta] φ [à cantora] φ
 2.i. [a aluna] φ [aceita] φ [cargo] φ [na empresa] φ
 2.ii. [a aluna] φ [aceitou] φ [cargo] φ [na empresa] φ
 2.iii. [a astrÓloga] φ [aceitou] φ [cargo] φ [na empresa] φ
 3.i. [a aluna] φ [apenas enviou] φ [flores] φ [à bailarina] φ
 3.ii. [a aluna] φ [apanhou sempre] φ [as melhores frutas] φ
 3.iii. [a astrÓloga] φ [apanhou sempre] φ [as melhores] φ
 4.i. [a aluna] *I*[após o conflito] *I*[repousou em seu quarto] *I*
 4.ii. [a aluna] *I*[apesar da discussão] *I*[ofereceu nova proposta] *I*
 4.iii. [a astrÓloga] *I*[apesar da discussão] *I*[ofereceu nova proposta] *I*

Em resumo, no conjunto de estruturas elaboradas, as variáveis ‘tonicidade das vogais’ e ‘distância entre os acentos’ foram controladas

de modo a assegurar a relação entre essas variáveis e as fronteiras prosódicas consideradas. A interação dessas variáveis, somada aos dois contextos segmentais controlados, a saber /a+a/ e /a+u/, resulta em 32 estruturas (4 fronteiras prosódicas X 2 contextos segmentais X 2 posições de acento na seqüência X 2 contextos de distância entre os acentos). Essas variáveis são esquematicamente apresentadas na tabela abaixo.

Tabela 1 – Variáveis controladas

Estrutura prosódica	Mesmo ϕ
	$\phi+\phi$ não-ramificado
	$\phi+\phi$ ramificado
Contexto segmental	I+I
	/a+a/ /a+u/
Posição do acento na seqüência vocálica	V+V'
	V'+V
Distância entre os acentos	1 sílaba
	2 sílabas

As sentenças elaboradas a partir dessas estruturas foram lidas duas vezes por três informantes representantes do dialeto paulista de São José do Rio Preto, totalizando 264 ocorrências (44x2x3). As sentenças foram gravadas digitalmente (em aparelho de MD) e analisadas por meio do programa PRAAT. A seguir, apresentamos os resultados obtidos.

4. EVIDÊNCIAS ENTOACIONAIS DA PROEMINÊNCIA DA FRASE FONOLÓGICA

Nesta seção, consideramos os contextos segmentais controlados para analisar o bloqueio da DG, a saber /a+á/ e /á+a/, e os contextos segmentais de bloqueio da EL, a saber /a+ú/ e /á+u/, para investigar como as informações entoacionais estão associadas ao elemento proeminente de ϕ . Em outras palavras, queremos obter evidências do embate que se estabelece no domínio de ϕ : de um lado, a tendência à otimização da cadeia silábica (por meio da aplicação de processos

segmentais) e, de outro lado, a tendência a preservar estrutura portadora da informação sintática relevante, que se manifesta pela proeminência rítmica e entoacional de ϕ . A análise desse embate, de natureza teleológica segundo Abaurre (1996), permite estabelecer relação entre a estrutura entoacional e a estrutura prosódica do PB.

Tomando como ponto de partida a distribuição de eventos tonais para as seqüências em que apenas a segunda vogal é tônica, apresentamos, inicialmente, os resultados em que a seqüência segmental caracteriza o contexto da EL. Por meio da tabela (2), são explicitados os eventos tonais mais freqüentes em relação à estrutura prosódica (coluna 1) e a distância entre os acentos (colunas 2 e 3).

Tabela 2. Distribuição de eventos tonais da seqüência /a+ú/

Estrutura prosódica	Distância de 1 σ entre os acentos	OC	Distância de 2 σ entre os acentos	OC
Mesmo ϕ	[aluna <u>útil</u>] σ LH* H* LH*	4 2	[astróloga <u>útil</u>] σ LH* LH*	6
ϕ - ϕ \tilde{n} -ram	[aluna] ϕ [usa] LH* L* H*	3 3	[astróloga] ϕ [usa] LH* L* H*	4 2
ϕ - ϕ ram.	[aluna] ϕ [usa sempre] LH* LH* I+I	6	[astróloga] ϕ [usa sempre] LH* LH*	6
	[aluna]I [útil mas desorganizada]I LH* L* H*	5 1	[astróloga]I [útil mas desorganizada]I LH* LH* L*	4 2

Onde: "OC" indica as ocorrências encontradas para cada contorno encontrado (em um total de 6 analisadas); CAIXA-ALTA indica os acentos; sublinhado indica os contextos sujeitos ao sândi.

Considerando a distribuição dos eventos tonais da tabela (2), constata-se que, em um mesmo ϕ , há indícios de que a distância entre os acentos é relevante para a configuração dos tons. Quando a distância entre os acentos é de uma sílaba (coluna 2), ocorre um tom simples, após o tom complexo inicial.¹¹ Quando a distância é de duas sílabas entre os acentos (coluna 3), a opção por um tom complexo

LH* inicial se mantém, mas aumenta a ocorrência de um tom complexo em seguida. Essa mesma relação entre tipo de eventos tonais e distância entre os acentos pode ser observada entre I_s .

Entre ϕ_s , observa-se que somente quando o segundo ϕ for ramificado, há a ausência de evento tonal associado à segunda vogal da seqüência sujeita à EL, e a ocorrência de LH* associado ao elemento cabeça do ϕ ramificado. Nota-se que essa configuração não depende da distância entre os acentos das palavras sujeitas ao sândi, como identificado dentro de ϕ , mas da distância entre os elementos cabeça de ϕ . Se comparada essa distribuição de tons com a que se verifica quando o contexto é entre ϕ_s não-ramificados, verifica-se que, independentemente da distância entre os acentos de w (coluna 2 *versus* coluna 3), a opção preferencial é por um tom simples associado ao elemento cabeça do segundo ϕ . Também se atesta um tom complexo LH* associado ao primeiro elemento acentuado de ϕ , que sempre corresponde à primeira sílaba acentuada de I não-final.¹²

Esse quadro de distribuição de eventos tonais traz à tona uma relação complexa entre a configuração dos contornos entoacionais e a preservação de certas informações da estrutura prosódica do PB. A análise da tabela (3) permite visualizar que o contorno entoacional se configura de maneira que fica assegurada a proeminência relativa da cabeça de ϕ . O contraste entre as estruturas (11.1) e (11.2) permite verificar que, para ocorrer um evento tonal complexo associado ao elemento proeminente de ϕ , faz-se necessária a distância de três sílabas entre as cabeças de ϕ . Portanto, a configuração entoacional encontrada depende da distância entre os elementos proeminentes de ϕ . Queremos propor que essa configuração entoacional é evidência da importância em preservar o elemento relativamente mais proeminente de ϕ . Outra evidência a favor dessa interpretação é a segmental: a elisão da vogal em (11.2.ii) leva a um *output* com três sílabas entre os acentos de ϕ , o que gera uma configuração da grade métrica semelhante a (11.2.i); por outro lado, em (11.2.i), o bloqueio da EL garante essa distância entre os acentos, preservando, assim, a proeminência relativa de ϕ .

(11) 1.i. [a aluna] ϕ [usa] ϕ [chapéu branco] ϕ [em público] ϕ

ϕ		*		*	
ω		*		*	
Σ		*		*	
σ	*	*	*	*	*
		T Γ *		T*	

1.ii. [a astróloga] ϕ [usa] ϕ [chapéu branco] ϕ [em público] ϕ

ϕ		*		*	
ω		*		*	
Σ		*		*	
σ	*	*	*	*	*
		T Γ *		T*	

2.i. [a aluna] ϕ [usa sempre] ϕ [chapéu branco] ϕ [em público] ϕ

ϕ		*		*	
ω		*	*	*	
Σ		*	*	*	
σ	*	*	*	*	*
		T Γ *		T*	

2.ii. [a astróloga] ϕ [usa sempre] ϕ [chapéu branco] ϕ [em público] ϕ

ϕ		*		*	
ω		*	*	*	
Σ		*	*	*	
σ	*	*	*	(*)	*
		T Γ *		T*	

Em resumo, ao considerarmos os contextos de /a+ú/, vemos que, para a ocorrência de um evento tonal complexo, é necessário que dentro de ϕ haja a distância de duas sílabas entre os acentos de ω , e que entre ϕ_s haja a distância de três sílabas entre os acentos de ϕ .

Um resultado semelhante se constata quando é levado em conta o segundo contexto em que apenas a segunda vogal é acentuada para a DG. Esse contexto se caracteriza pelo bloqueio da DG, quando o acento da segunda vogal é interpretado como acento

de ϕ . Considerando a distribuição dos eventos tonais da tabela (3), verifica-se que em todos os contextos sempre houve um evento tonal associado ao elemento proeminente de ϕ . Em um mesmo ϕ , a sílaba interpretada como cabeça de ϕ (que também é candidata à DG) está preferencialmente associada a um tom simples. Entre ϕ_s , quando o segundo não é ramificado, também preferencialmente é associado um evento tonal simples à segunda sílaba acentuada que é cabeça de ϕ .

Tabela 3 – Distribuição de eventos tonais da seqüência /a+á/

Estrutura prosódica	Distância de 1 σ entre os acentos	OC	Distância de 2 σ entre os acentos	OC
Mesmo ϕ	[a aluna árabe] ϕ		[a astróloga árabe] ϕ	
	LH* L*	5	LH* L*	4
	L* LH*	1	LH* LH*	2
ϕ - ϕ não-ramificado	[a aluna] ϕ [age]		[a astróloga] ϕ [age]	
	LH* L*	4	LH* L*	6
	LH* H*	1		
L* LH*	1			
ϕ - ϕ ramif.	[a aluna] ϕ [age sempre]		[a astróloga] ϕ [age sempre]	
	LH* LH*	4	LH* LH*	4
	L* LH*	2	L* LH*	2
I+I	[a aluna] I		[a astróloga] I	
	[ávida por justiça]		[ávida por justiça]	
	LH* L*	5	LH* L*	6
H* L*	1			

Onde: “OC” indica as ocorrências encontradas para cada contorno encontrado (em um total de 6 analisadas); CAIXA-ALTA indica os acentos; sublinhado indica os contextos sujeitos ao sândi.

Nota-se que em nenhum desses casos parece ter ocorrido a DG, embora no contexto ‘*aluna age*’ tenha sido encontrada a maior variedade de eventos tonais associados às cabeças de ϕ e uma alta variação da duração do intervalo vocálico. Essas variações, entretanto, não estão dissociadas, uma vez que (i) quando se configura apenas um evento tonal LH* associado à cabeça do segundo ϕ , encontra-se um aumento da duração do intervalo vocálico compreendido pelas vogais iguais /a+á/; (ii) quando se associa um evento tonal H* ao segundo ϕ , a medida de duração do intervalo vocálico /a+á/ é a mais

baixa. Esses resultados são indícios de haver uma tensão que não se verifica na maioria das ocorrências registradas. Quando há ‘espaço’ suficiente, os eventos tonais LH* e L* são associados aos elementos cabeça de ϕ . Essa tensão parece decorrer da proximidade dos acentos de ϕ : por um lado, é preciso assegurar a proeminência de ϕ , por outro, há pouco espaço entre essas proeminências.

Ainda considerando a distribuição dos eventos tonais da tabela (3), constata-se que somente entre $\phi+\phi$ ramificado a segunda vogal candidata ao sândi não está associada a um evento tonal. Nesse contexto, o evento tonal ocorre associado ao elemento mais proeminente de ϕ , o qual não coincide com a vogal candidata à DG, e o processo ocorre. Encontra-se, também, a predominância de um tom complexo LH* associado ao elemento proeminente de ϕ . Essa regularidade decorre da distância de três sílabas que há entre os elementos proeminentes de ϕ . Nota-se que, para essa mesma estrutura prosódica, foi encontrada a mesma configuração entoacional, quando considerado outro contexto segmental (cf. tabela 2).

Por fim, entre I_s há uma grande regularidade dos tipos de eventos tonais e, em todas as ocorrências analisadas, a segunda vogal acentuada está associada preferencialmente ao tom L*, independentemente de haver uma ou duas sílabas de distância entre as sílabas acentuadas ou mesmo naquelas em que a duração permite afirmar que há DG. Esse resultado mostra que, diferentemente do que foi identificado no contexto característico da EL, a DG não é bloqueada quando a sílaba acentuada estiver em fronteira de I . Essa diferença entre o bloqueio da EL e da DG é indício de que as restrições que operam em PB dependem do tipo do processo envolvido e da estrutura prosódica que as contém. Contudo, um fato comum salta aos olhos: a importância da proeminência de ϕ .

A importância desse domínio em PB pode ser atestada também nos contextos em que a primeira vogal da seqüência é acentuada. Ao considerarmos os dados da tabela (4), a seguir, verifica-se que sempre ocorre um evento tonal LH* associado à primeira sílaba da seqüência candidata à DG, à exceção de algumas ocorrências dentro de ϕ . No entanto, o processo não é bloqueado por esse evento tonal que, em todas as estruturas, é a primeira sílaba acentuada do domínio I .

Tabela 4. Distribuição dos eventos tonais da seqüência /á+a/

Estrutura prosódica	Distância de 1 σ entre os acentos	OC	Distância de 2 σ entre os acentos	OC
Mesmo ϕ	[o marajá afoito] ϕ LH* L*	3	[o marajá africano] ϕ LH* L*	4
		3	L* LH*	2
ϕ - ϕ não-ramificado	[o marajá] ([aceita]) LH* L*	5	[o marajá] ([apanhou]) LH* L*	5
		1	H*	1
ϕ - ϕ ramif.	[o marajá] ϕ [apenas aceitou] LH* LH*	6	[o marajá] ϕ [apanhou sempre] LH* LH*	6
	I+I		[o marajá] I [apesar da confusão] LH* L*	6

Onde: “OC” indica as ocorrências encontradas para cada contorno encontrado (em um total de 6 analisadas); CAIXA-ALTA indica os acentos; sublinhado indica os contextos sujeitos ao sândi.

O resultado descrito é uma evidência entoacional a favor da interpretação de que a reestruturação silábica desencadeada pela DG está sujeita ao princípio proposto por Bisol (1996a). Segundo esse princípio, a aplicação de processos segmentais de juntura opera da esquerda para a direita e, portanto, apenas os elementos proeminentes que estão à direita geram o bloqueio desses processos. Como em /á+a/ a sílaba acentuada portadora de evento tonal ocorre na posição mais à esquerda na fronteira vocabular, o bloqueio não se verifica.

Cabe salientar que a direção da reestruturação silábica, visível por meio da ocorrência da DG quando /v'+v/, dialoga com o fato de PB ser uma língua de recursividade sintática à direita (cf. CINQUE, 1993). Abaurre (1996) já havia explorado a relação entre o bloqueio de sândi vocálico e a necessidade de preservação da proeminência de f, que traz informação sobre a direção da recursividade da língua quando há bloqueio da DG e da EL. Portanto, fica claro que não é o fato de haver um evento tonal que gera o bloqueio de processos segmentais, mas o fato de este evento carregar uma informação sobre a direção da recursividade sintática do PB.

A segunda seqüência vocálica estudada em que a primeira vogal da seqüência é acentuada trata-se da EL. Esse processo é sempre bloqueado em todos os contextos independentemente da distância entre os acentos. Assegurados os mesmos contextos rítmicos e prosódicos que os controlados para o contexto de DG, verifica-se que a EL tem comportamento diferente: o acento impede que a vogal /a/ seja eliminada. O encontro de núcleos silábicos é resolvido por meio da ditongação. Mais uma vez, constata-se que os processos de sândi vocálico têm comportamento diferente em PB.

Embora haja diferença no plano segmental, são várias as semelhanças na distribuição de eventos tonais quando comparadas as mesmas estruturas prosódicas. Por meio da tabela (5), constata-se que sempre ocorre um evento tonal LH* associado à primeira sílaba da seqüência. Esse evento tonal está associado à primeira sílaba que carrega acento primário da sentença, que é simultaneamente a cabeça de ϕ e o primeiro elemento proeminente dentro de *I* e de *U*. Mais uma vez, o contraste entre as estruturas em que está controlada a cabeça de ϕ possibilita observar que a distância de três sílabas entre os elementos proeminentes de ϕ é suficiente para ocorrer um evento tonal complexo. Não havendo essa distância, o evento tonal passa a ser um tom simples.

Tabela 5 – Distribuição dos eventos tonais da seqüência /á+u/

Estrutura prosódica	Distância de 2 σ	Realização
Mesmo ϕ	[o marajá humilhado] ϕ LH* L*	o maraj[au]milhado * o maraj[ul]milhado
ϕ - ϕ não-ramificado	[o marajá] ϕ [usará] LH* L*	o maraj[au]sará * o maraj[ul]sara
ϕ - ϕ ramificado	[o marajá] ϕ [usará sempre] LH* LH*	o maraj[au]sará sempre * o maraj[ul]sará sempre
I+I	[o marajá] <i>I</i> [humilhado pela cantora] LH* L*	o maraj[au]milhado pela cantora] * o maraj[ul]milhado pela cantora

Onde: CAIXA-ALTA indica os acentos; sublinhado indica os contextos sujeitos ao sândi. Em todas as ocorrências, a realização do contorno entoacional foi a mesma.

Por fim, faz-se necessário comparar as quatro tabelas em que é apresentada a distribuição dos eventos tonais para as seqüências de vogais átonas seguidas de tônicas e das seqüências de vogais tônicas seguidas de átonas. Duas regularidades devem ser destacadas:

- (i) a presença de evento tonal associado ao elemento proeminente de ϕ ;
- (ii) a preferência pela alternância baixo-alto entre os eventos tonais, que, freqüentemente, se configura em LH* L*, quando a distância é de até duas sílabas átonas entre os acentos de ϕ , ou em LH* LH*, quando a distância é igual a três sílabas entre os acentos.

O fato expresso em (i) atesta a relevância da proeminência do domínio ϕ para a distribuição dos eventos tonais em PB. Já a constatação feita em (ii) revela uma característica do contorno entoacional do PB como sendo constituído com base nos elementos mais proeminentes de ϕ . Para garantir essa proeminência, há uma diferença de altura entre os tons de modo que o tom complexo LH* é seguido de um tom simples L* quando há proximidade entre os acentos de ϕ . Daí atesta-se a relevância da distância entre os elementos proeminentes de ϕ para a configuração entoacional. Nota-se que a língua, valendo-se dessa característica, minimiza um possível efeito de quebra de euritmia que a proximidade dos acentos geraria na medida em que ou atribui tons de altura diferentes a cada elemento proeminente de ϕ , quando esses estão próximos, ou não atribui tons a um dos elementos da estrutura em jogo, quando a sílaba acentuada não carrega o elemento mais proeminente de ϕ , não veiculando, portanto, informação sobre a estrutura sintática da língua.

A partir desse quadro, a pergunta a fazer é: essa distribuição dos eventos tonais consiste em uma 'estratégia entoacional' para resolver o choque de acentos que decorre da aplicação de processos de reestruturação silábica? Ou ainda, os tons são alterados em sua configuração para assegurar a euritmia? A resposta pode ser dada se levada em conta a distribuição dos eventos tonais em contextos em que as vogais sujeitas ao sândi vocálico não carregam acento e o número de sílabas átonas não gera o choque de acentos dentro de ϕ e entre ϕ .

Em busca dessa resposta elaboramos um conjunto de sentenças em que a distância entre os acentos é de duas, três e quatro sílabas, respectivamente, nos quatro contextos prosódicos que temos considerado. Em todas as estruturas, há o contexto de sândi entre as fronteiras prosódicas relevantes, de modo que a distância entre os acentos passa a ser de uma, duas e três sílabas, respectivamente. A distribuição dos eventos tonais associados às sílabas tônicas é esquematicamente apresentada na tabela (6).

Tabela 6 – Distribuição dos eventos tonais e a distância entre os acentos

Estrutura Prosódica	Distância entre os acentos		
	2σ	3σ	4σ
Mesmo φ	[a aLUna aFOIta] φ LH* L*	[a aLUna afriCAna] φ LH* L*	[a asTRÓloga afriCAna] φ LH* L*
φ-φñ-ram.	[a aLUna] φ [aCEItA] LH* L*	[a aLUna] ([aceiTOU] LH* L*	[a asTRÓloga] φ [aceiTOU] LH* L*
φ-φram.	[a aLUna] φ [aPEnas enviOU] LH* H* L*	[a aLUna] φ [apaNHOUs SEMpre] LH*	[a asTRÓloga] φ [apaNHOUs sempre] LH*
I+I	[a aLUna] I[aPÓS o conflito] LH* L*	[a aLUna] I[apeSAR da discussão] LH* L*	[a asTRÓloga] I[apesAR da discussão] LH* L*

Onde: CAIXA-ALTA indica os acentos; sublinhado indica os contextos sujeitos ao sândi. Em todas as ocorrências, a realização do contorno entoacional foi a mesma.

A configuração entoacional apresentada na tabela acima revela que a distribuição dos eventos tonais não é alterada em função da proximidade dos acentos. Independentemente do número de sílabas átonas entre os acentos, encontra-se um tom complexo LH* associado à primeira sílaba tônica seguido de um tom simples L* na maioria dos contextos prosódicos, ou um tom LH* somente quando o segundo φ é ramificado.

Comparando-se todos os quadros de distribuição dos eventos tonais, conclui-se que a variação desses eventos está relacionada aos domínios prosódicos e não é pertinente afirmar que haja uma

‘estratégia entoacional’ para resolver o choque de acentos que decorre da reestruturação silábica. Desse modo, propomos que em contexto de choque de acentos a estratégia entoacional operante em PB consiste basicamente em otimizar a alternância dos tons que necessariamente estão associados aos elementos proeminentes de f. Em outras palavras, as diferenças de altura entre os tons têm como consequência impedir que certas seqüências possam ser percebidas como perturbadoras da eurritmia.

Essa interpretação encontra respaldo na observação de Abousalh (1997) sobre a relação entre contornos entoacionais e resolução de choque de acentos em PB. Quando discute a influência da fronteira de ϕ na resolução dos choques de acentos, a pesquisadora brasileira afirma, após análise de um exemplo de choque de acentos, que

as diferenças tonais observadas entre os acentos dos choques parecem ser um reflexo mais geral de marcação melódica de final de enunciado, que não é exclusivo de seqüências que contêm *clashes*. (p. 125)

Também quando discute a influência da fronteira de *I*, Abousalh (1997) argumenta que

o que nos parece mais apropriado dizer é que, a partir do momento em que ocorre uma fronteira de *I* entre os acentos de um choque, ela serve para desfazer a sensação de colisão acentual. (p. 117)

Dessas afirmações e dos resultados que encontramos, fica claro que a distribuição dos eventos tonais em PB não está sujeita a restrições rítmicas, ou seja, o efeito de choque de acentos não leva a mudanças no contorno entoacional que possam ser tomadas como uma estratégia disponível para resolver a colisão entre os acentos.

Concluimos esta seção tendo encontrado evidências de que a distribuição dos eventos tonais está vinculada à constituição dos domínios prosódicos, especificamente ao domínio ϕ , e o efeito de choque de acentos desencadeado pelo sândi vocálico pode ser minimizado nos diferentes domínios. Isso pode ser observado inclusive

dentro de ϕ pelo fato de os tons associados às proeminências relativamente mais fortes de ϕ serem, preferencialmente, de alturas diferentes.

5. CONSIDERAÇÕES FINAIS

Neste artigo, o nosso objetivo foi o de apresentar resultados de uma análise entoacional de contextos sujeitos à degeminação e à elisão que fornecessem subsídios sobre como se dá em PB a interpretação fonológica da informação da recursividade sintática.

O *corpus* de sentenças foi elaborado tendo sob controle o contexto segmental, a posição do acento na seqüência vocálica, a distância entre os acentos considerados e as fronteiras prosódicas em que ocorrem os contextos de sândi analisados. A análise entoacional seguiu a abordagem proposta por Ladd (1996) e a hierarquia prosódica considerada, a proposta por Nespor & Vogel (1986).

A análise de como se dá a associação dos eventos tonais à cadeia segmental, quando são consideradas seqüências de vogais átonas seguidas de tônicas e de vogais tônicas seguidas de átonas, revela:

- (i) a presença de evento tonal associado ao elemento proeminente de ϕ ;
- (ii) a preferência pela alternância baixo-alto entre os eventos tonais, que freqüentemente se configura em LH* L*, quando a distância é de até duas sílabas átonas entre os acentos de ϕ , ou em LH* LH*, quando a distância é igual a três sílabas entre os acentos.

Essas regularidades da configuração entoacional são atestadas não somente quando está em conflito a aplicação do processo de sândi e o choque acentual, mas também quando a colisão acentual não se verifica. Em outras palavras, foram encontradas evidências de que a configuração entoacional está relacionada à constituição dos domínios prosódicos, sendo a proeminência de ϕ uma informação fundamental para a distribuição dos eventos tonais em PB.

Esse fato de o acento nuclear de ϕ ser portador de evento tonal, somado à asserção de que esse mesmo acento de ϕ indica a direção

de encaixamento sintático da língua, é tomado como evidência de que a fonologia interpreta a informação sintática e, assim, se estabelece o domínio prosódico ϕ , cuja proeminência é relevante quer para o ritmo, quer para a entoação do PB. Ou seja, a distribuição de eventos tonais que identificamos é uma pista do modo pelo qual a fonologia traduz a informação sobre a direção do encaixamento sintático, uma pista de como se dá a interface sintaxe-fonologia.

NOTAS

* Texto apresentado na mesa-redonda “A fonologia na interface com outros componentes da gramática” durante o *I Congresso Internacional de Fonética e Fonologia & VII Congresso Nacional de Fonética e Fonologia* – Belo Horizonte – UFMG.

¹ Mantivemos os mesmos exemplos de Bisol (1993).

² Os primeiros modelos de análise da relação entre fonologia e demais módulos da gramática são elaborados por Selkirk (1984) e Nespor & Vogel (1986). Selkirk (1984) propõe que a fala seja organizada hierarquicamente em cinco constituintes prosódicos, a saber: sílaba, pé, palavra fonológica, frase fonológica e frase entoacional. Já em Nespor & Vogel (1986) são propostos sete constituintes, a saber: sílaba, pé, palavra fonológica, grupo clítico, frase fonológica, frase entoacional e enunciado fonológico.

³ Cabe lembrar que o algoritmo de f utilizado por Abaurre (1996) é aquele proposto por Nespor & Vogel (1986). Cf. na terceira seção o algoritmo de formação de ϕ .

⁴ A relação entre esses dois fatores é estabelecida, segundo essa abordagem teórica, nos seguintes termos: “*rhythm is represented (...) in terms of the grid, which is constructed on the basis of information contained in the prosodic tree. That is, each syllable is assigned one position, marked by an x on the first grid level. Subsequently, the DTE (designated terminal element) of each higher prosodic category (excluding [...] the clitic group), is assigned one additional x on a separate grid level*” (NESPOR, 1990: 245).

⁵ O bloqueio da regra de retração de acentos está submetido, segundo esses autores, a certas restrições relacionadas ao estabelecimento das fronteiras de ϕ . Dessas restrições tratamos em Tenani (2002).

⁶ Em Tenani (2002) também são consideradas medidas de duração na análise das estratégias de resolução de choque acentual.

⁷ Selkirk (1984) dedica todo um capítulo para argumentar que *‘the tonal elements making up the pitch contour of the intonational phrase are assigned directly (and freely) to surface syntactic structure, and that it is on the basis of this assignment that the essentially semantic properties of the focus of the sentence are defined’* (p. 24).

⁸ Um exemplo dessa afirmação se verifica na conclusão de Beckman & Pierrehumbert (1986) sobre as características entoacionais do Inglês e do Japonês. As autoras afirmam que ambas as línguas *‘organize the tone features into a hierarchy of prosodic structures, from the grouping of tones into pitch accents at local level to the choice of phrase-terminal tones and the manipulation of pitch range over larger domains’* (p. 305).

⁹ Em PB, essa fronteira de *I* não bloqueia a DG e a EL, quando ambas as vogais são átonas (cf. TENANI, 2002), mas resta saber se *I* é importante para o bloqueio da DG e da EL em PB, quando uma das vogais é acentuada. Em Português Europeu, a fronteira de *I* bloqueia a aplicação de sândi vocálico quando ambas as vogais são átonas, mas não atua quando uma das vogais é tônica (cf. FROTA, 1998).

¹⁰ Cabe remeter o leitor à conclusão de Massini-Cagliari (1992: 38) sobre os correlatos do acento: (i) “no nível lexical, os principais correlatos do acento são (em ordem decrescente de importância): duração, intensidade e qualidade vocálica”; (ii) “no nível frasal, o acento do português é caracterizado por uma variação do padrão entoacional que se sobrepõe a uma sílaba tônica em nível lexical. Em outras palavras, para que uma sílaba tônica em nível lexical o seja em nível frasal é preciso que ela ocorra num determinado momento do padrão entoacional (relativo a todo o enunciado) em que esteja havendo variações relativamente grandes de F0 em relação a todo o resto do enunciado”.

¹¹ A outra possibilidade que se apresenta nesse contexto prosódico é apenas a ocorrência de um evento tonal associado à cabeça de *f*, ficando sem tom a primeira sílaba da sentença.

¹² Encontra-se em Tenani (2002) uma descrição das características entoacionais de *I* no PB.

REFERÊNCIAS BIBLIOGRÁFICAS

ABAURRE, M. B. M. Acento frasal e processos fonológicos segmentais. *Letras de Hoje*, v.2, n. 31, p. 41-50, 1996.

ABOUSALH, E. F. *Resolução de choques de acento no português brasileiro*. Dissertação. 1997. (Mestrado) – Unicamp, Campinas.

- BECKMAN, M.; PIERREHUMBERT, J. Intonational structure in Japanese and English. *Phonology Yearbook*, n. 3, p. 255-310, 1986.
- BISOL, L. A elisão, uma regra variável. *Letras de Hoje*, v. 1, n. 35, p. 319-330, 2000.
- BISOL, L. Sândi externo: o processo e a variação. *Gramática do Português Falado*, Campinas: Editora da Unicamp, v. 5, p. 55-96, 1996a.
- BISOL, L. O sândi e a ressilabação. *Letras de Hoje*, v. 2, n. 31, p. 159-168, 1996b.
- BISOL, L. Sândi vocálico externo. *Gramática do Português Falado*. Campinas: Editora da Unicamp, v. 2, p. 21-38. 1993.
- CINQUE, G. A null theory of phrases and compound stress. *Linguistic Inquiry*, v. 2, n. 24, p. 239-297, 1993.
- FROTA, S. Nuclear falls and rise in European Portuguese: a phonological analysis of declarative and question intonation. 2001. Dissertação (Mestrado) – Universidade de Lisboa, Lisboa.
- FROTA, S. *Prosody and focus in European Portuguese*. 1998. Tese (Doutorado) – Universidade de Lisboa, Lisboa. Publicado por Garland Publishing (series Outstanding Dissertations on Linguistics). New York/London, 2000.
- GUIMARÃES, M. Repensando a interface sintaxe-fonologia a partir do axioma de correspondência linear. 1998. (Dissertação (Mestrado) – Unicamp, Campinas.
- HAYES, B. The prosodic hierarchy in meter. In: KIPARSKY, P.; YOUMANS, G. (Ed.). *Rhythm and meter. Phonetics and phonology 1*. New York: Academic Press, 1991. p. 201-260.
- HAYES, B.; LAHIRI, A. Bengali intonational phonology. *Natural Language & Linguistic Theory*, v. 1, n. 9, p. 47-96, 1991.
- INKELAS, S.; ZEC, D. Syntax-Phonology interface. In: GOLDSMITH, J. (Ed.). *The handbook of phonological theory*. Oxford: Blackwell, 1995. p. 535-549.
- INKELAS, S.; ZEC, D. *The phonology-syntax connection*. Chicago: University of Chicago Press, 1990.
- LADD, D. R. *Intonational Phonology*. Cambridge: Cambridge University Press, 1996.
- LADD, D. R. An introduction to intonational phonology. In: DOCHERTY, G.; LADD, D. R. (Ed.). *Papers in laboratory phonology 2*. Cambridge: Cambridge University Press, 1992. p. 321-334.
- MORAES, J. A. Correlats acoustiques de l'accent de mot em Portugais Brésilien. In: *XI International Congress of Phonetic Sciences*, Tallinn: Estônia, v. 3, p. 313-316, 1987.

- MASSINI-CAGLIARI, G. *Acento e ritmo*. São Paulo: Contexto, 1992.
- NESPOR, M. Setting parameters at a prelexical stage. ms, 1994.
- NESPOR, M.; VOGEL, I. *Prosodic Phonology*. Dordrecht-Holland: Foris Publications, 1986.
- PIERREHUMBERT, J. The phonology and phonetics of English Intonation. 1980. Dissertation (Ph.D.) – The MIT Press, Cambridge.
- PIERREHUMBERT, J.; BECKMAN, M. *Japanese tone structure*. Cambridge: The MIT Press, 1988.
- SANDALO, F.; TRUNCKENBRODT, H. Some notes on phonological phrasing in Brazilian Portuguese. *The MIT Working papers 42*. Cambridge: The MIT Press, 2002.
- SELKIRK, E. O. *Phonology and syntax, the relation between sound and structure*. Cambridge: Cambridge University Press, 1984.
- TENANI, L. E. Domínios prosódicos no Português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos. 2002. Tese (Doutorado) – Unicamp, Campinas.

Fonologia prosódica e teoria da otimalidade: reflexões sobre a interface sintaxe e fonologia na formação de sintagmas fonológicos¹

Filomena Sandalo
UNICAMP

Abstract

This paper discusses phonological phrasing in Brazilian Portuguese. The data suggest that phonological phrasing can access a syntactic relation such as SPEC-HEAD agreement in this language.

1. INTRODUÇÃO

O tema deste estudo é a relação entre sintaxe e fonologia na formação de sintagmas fonológicos no português brasileiro (PB). A perspectiva de estudo é a teoria da otimalidade aplicada ao estudo da fonologia prosódica, tendo como ponto de partida a proposta de Truckenbrodt (1999). Não se trata, entretanto, de meramente aplicar a esta teoria para se obter uma modelagem nova de velhos problemas já tratados e entendidos a partir de outros modelos gerativos. Busca-se aqui trazer para a discussão fenômenos fônicos pouco discutidos dentro do quadro gerativo e argumentar em favor de um modelo representacional. Em outras palavras, busca-se empregar a teoria da otimalidade para uma discussão mais abrangente e acurada de aspectos prosódicos do português e de outras línguas. Como um exemplo de fenômenos pouco discutidos, podemos citar a entoação. A entoação tem sido freqüentemente excluída de trabalhos em fonologia gerativa do português brasileiro. Frota & Vigário (2000) é um dos raros estudos sobre entoação no português brasileiro a partir da teoria gerativa. Frota & Vigário têm como base o trabalho de Frota (1998), que associa o estudo sobre fonética da entoação de Pierrehumbert (1980) e de Ladd (1996) à fonologia prosódica de Nespor & Vogel (1986). A premissa principal dessa fonologia é que a corrente fônica é organizada hierarquicamente em domínios; o suporte tradicional para se acreditar nesses domínios hierárquicos tem sido os domínios de aplicações de certos processos fonológicos segmentais como sândi, mas a entoação também tem sido considerada importante em alguns trabalhos recentes, como Hayes & Lahiri (1991) e Truckenbrodt (1995). Segundo Frota, o domínio da entoação no português europeu (PE) é o sintagma intonacional (I). Frota & Vigário afirmam que o domínio da entoação no PB é o sintagma fonológico (ϕ). Segundo Frota & Vigário (2000: 14, 17):

“Se considerarmos apenas os fs em posição não inicial e não final de I, os resultados de presença de acento tonal são de 80% para o PB contra apenas 27% para o PE. A exclusão da posição inicial e final de I é importante dado que os fs nessas posições são portadores de acento tonal por razões independentes: o f final porque constitui o elemento mais proeminente de I e é, portanto, o portador do acento tonal nuclear nas duas variedades; o f inicial porque, no PE, pode conter um acento tonal cuja função é assinalar o início de I. Em síntese, f é um domínio entoacionalmente robusto no PB, mas não no PE. No PE, tal robustez entoacional é apenas caracterizadora do domínio I, como aliás já tinha sido notado em trabalhos anteriores.” (cf. FROTA, 1998)

Há duas visões principais sobre fonologia prosódica no quadro da fonologia gerativa derivacional: a abordagem rotulada de *end-based* (SELKIRK, 1986) e a abordagem rotulada de *relation based* (NESPOR & VOGEL 1986, HALE & SELKIRK 1987). A diferença principal entre estas duas perspectivas está na natureza da informação sintática que pode ser tomada pelos algoritmos de formação dos domínios prosódicos. Segundo a primeira abordagem, as condições de mapeamento em domínios prosódicos somente podem acessar fronteiras sintáticas. Segundo a outra abordagem, as condições de mapeamento em domínios prosódicos podem acessar a relação entre núcleos e complementos (e certos constituintes adjacentes). Frota & Vigário (2000) assumem a perspectiva de Nespor & Vogel. Segundo esta perspectiva, a formação de fs se dá a partir do seguinte algoritmo:

(1) ϕ no Português adaptado de Frota & Vigário (2000: 13)

Todas as cabeças lexicais e os elementos à sua esquerda até à cabeça lexical precedente constituem um ϕ .

Um ϕ deve ser constituído por mais do que uma palavra fonológica, permitindo que fs sejam opcionalmente reestruturados, formando um único ϕ com um complemento não ramificado.

Nespor & Vogel (1986) propõem que sintagmas fonológicos sejam formados por um algoritmo de duas partes. Primeiro, sintagmas fonológicos são construídos fazendo-se referência aos núcleos lexicais. Pode-se notar no algoritmo acima que cada núcleo lexical implica estabelecer uma fronteira prosódica, no caso uma

fronteira de f. Em uma segunda etapa, o algoritmo permite que um núcleo e seu complemento possam se reestruturar em um único sintagma fonológico. Esta reestruturação está sujeita à condição de que o complemento não pode ser ramificado.

A previsão, portanto, para a entoação do PB, ao se adotar o algoritmo de Nespor & Vogel, é que haja a seguinte variação livre onde cada domínio representado por f carregue um acento frasal:

2. (Café ϕ) (quente ϕ) (queima ϕ) ~ (Café quente ϕ) (queima ϕ)
3. (Café ϕ) (quente ϕ) (queima ϕ) (a boca ϕ) ~ (Café quente ϕ) (queima ϕ) (a boca ϕ)
~ (Café quente ϕ) (queima a boca ϕ)

Há um problema empírico a ser enfrentado no estudo de entoação no PB, entretanto, ao se adotar a perspectiva de Nespor & Vogel, a variação livre prevista não parece ocorrer. Um estudo realizado por Sandalo & Truckenbrodt (2002), a partir de dados de seis falantes nativos do PB da região sudeste, sugere que uma sentença com o padrão sintático N(ome)A(djetivo)V(erbo), quando lida como manchete de jornal ou quando dada como resposta a uma pergunta do tipo *o que aconteceu?* (isto é, quando a sentença como um todo é informação nova), apresenta o padrão entoacional em (4), enquanto uma sentença com o padrão sintático N A V N apresenta o padrão em (5).²

- (4) (Café quente)_p(queima a boca)_p

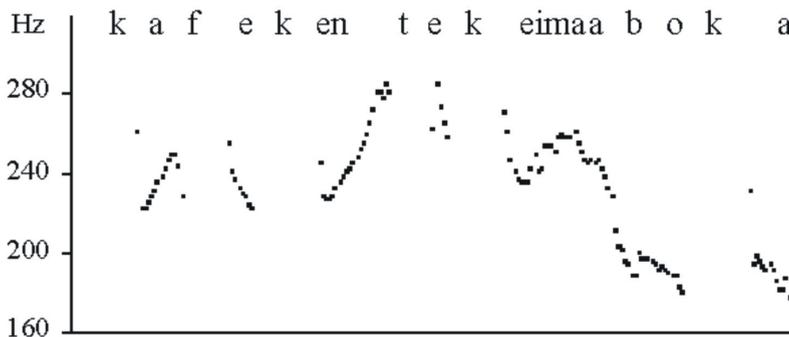


Figura 1: Entoação NAVN

(5) (café)_p(quente)_p(queima)_p

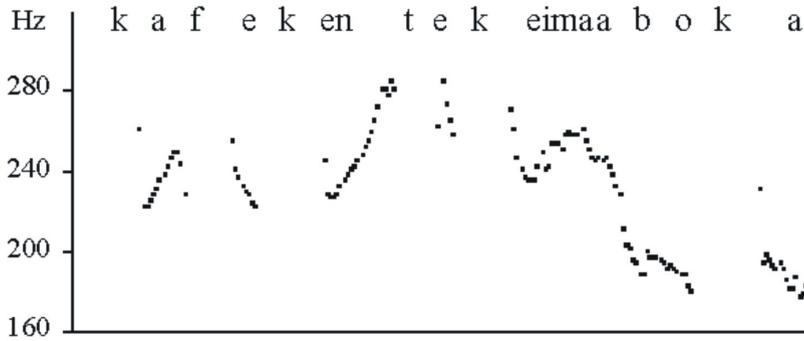


Figura 2: Entoação NAV

As curvas entoacionais observadas no estudo de Sandalo & Truckenbrodt sugerem que (i) cada acento lexical em posição não final é associado a uma subida na entoação e (ii) a distinção entre acento lexical e acento frasal é refletida no tamanho desta subida. Isto é, o acento frasal é mais alto. Há apenas um acento frasal em *café quente* na sentença em 5 (*café quente queima a boca*), ou seja, o sintagma sintático equivale a um sintagma fonológico. Mas na sentença em 4, *café quente queima*, o tom alto em *café* é tão alto quanto o de *quente*. Isto é, há dois acentos frasais, ou seja, dois sintagmas fonológicos no sintagma nominal *café quente*. Neste estudo preliminar, uma realização de N A V, em que N A comporta um único acento frasal, e uma realização de N A V N, em que cada núcleo lexical carrega um acento frasal, somente foram atestadas no caso de topicalização de um dos constituintes sintáticos no primeiro caso (que muitas vezes vem associada a uma pausa ou por certos alongamentos vocálicos) ou em fala lenta ou enfatizada no segundo caso, não sendo, portanto, opções não-marcadas. Parece haver uma relação entre o tamanho do sintagma fonológico e a transitividade do verbo.

Outro aspecto prosódico pouco entendido no quadro gerativo com base no português e nas línguas ameríndias, embora também tenha como domínio o sintagma fonológico em várias línguas (cf.

SANDALO & TRUCKENBRODT 2002), é o fenômeno conhecido como retração acentual em caso de colisão de acentos. O fenômeno foi notado para o PB por Major (1985) e discutido em mais detalhe por Abousalh (1997). Abousalh, com base em dados de produção, afirma que, embora casos de colisão acentual com retração de acento não sejam abundantes em seus dados, foi possível notar que colisões acentuais nunca foram desfeitas na fronteira de sintagmas fonológicos. Além disso, a autora nota que os únicos casos de retração acentual que ocorrerem em seu *corpus* estavam dentro de um sintagma fonológico. Abousalh também emprega o algoritmo de Nespor & Vogel para seu estudo e, dado o número reduzido de retrações de acento encontrado em seu *corpus*, a autora nada pode acrescentar sobre a possibilidade de variação livre. Sandalo & Truckenbrodt (2002) revisitam o fenômeno com base em um *corpus* de percepção. Os mesmos domínios foram notados, quando comparados aos domínios notados através dos dados de entoação. Isto é, notamos que os fatos sobre retração de acento questionam fortemente a variação livre prevista pelo algoritmo de Nespor & Vogel. O estudo de Sandalo & Truckenbrodt tem como base principal de reflexão a percepção de retração acentual em PB, já que é tradição na literatura fonológica gerativa estabelecer fronteiras prosódicas com base em dados de percepção. Muitas línguas foram estudadas a partir da percepção de desfazimento de colisão acentual através de retração de acento, como, por exemplo, Nespor & Vogel 1979, Hayes 1984, Hayes & Puppel 1985, Visch 1989, Nespor & Vogel 1989, Gussenhoven 1991, Vogel et al. 1995.³

O presente estudo defende com Frota & Vigário (2000) e com Abousalh (1997) que o domínio da entoação e do desfazimento de colisão acentual no PB é o sintagma fonológico, dada a relação próxima desses domínios com sintagmas sintáticos. Este estudo, entretanto, procura mostrar que é necessário desenvolver um novo modelo para a formação de sintagmas fonológicos que seja capaz de prever variações livres quando elas realmente ocorrerem, bem como de prever a inexistência de variação livre quando elas não existirem.

A possibilidade de inexistência da variação livre como prevista pela teoria de Nespor & Vogel tem sido largamente ignorada pela literatura fonológica, sendo uma exceção Ghini (1993) em um estudo sobre o italiano. Ghini (1993) defende uma análise em termos de fatores de euritmia para o desenvolvimento de um algoritmo de formação de sintagmas fonológicos do italiano, que parecem ser necessários também para o português (cf. SANDALO & TRUCKENBRODT 2002). É possível que a falta de discussão sobre a questão da ausência de verdadeira variação livre na formação de sintagmas fonológicos seja devida ao fato de que certos padrões prosódicos somente podem ser entendidos a partir de uma preocupação com a configuração da gramática como um todo, o que acaba se tornando menos freqüente em uma tradição de estudo dividida em módulos independentes de análise. Espera-se poder mostrar que um processamento paralelo da fonologia e da sintaxe é necessário para se entender pelo menos alguns aspectos da prosódia de línguas naturais. Uma previsão importante desta linha de pesquisa é que a fonologia pode não ter um papel meramente interpretativo (i.e., acessar apenas informações sintáticas bastante restritas), como se costuma assumir no quadro da teoria gerativa derivacional. Este estudo visa mostrar que esta previsão está correta.

A seção 2 apresenta o *corpus* deste estudo, e a seção 3 desenvolve a análise dos dados.

2. CORPUS

O *corpus* para este estudo foi elaborado através de um questionário testando aceitabilidade para retração de acento realizado com seis falantes do português brasileiro. Este *corpus* inclui por volta de 150 sentenças que manipulam diversos tipos de estruturas sintáticas. Uma leitura de informação nova (foco largo) das sentenças elicitadas foi garantida, pedindo-se que os falantes consultados julgassem se uma retração de acento, em situação de colisão acentual, era ou não permitida no caso de as sentenças serem lidas como manchetes de jornal ou como resposta para uma pergunta do tipo *o que*

aconteceu?. Garantir foco largo das sentenças elicitadas é importante já que foco estreito ou ênfase bloqueia retração de acento, e topicalização de informação dada força retração a ocorrer onde ela normalmente não ocorre. Três dos falantes consultados foram entrevistados por mim pessoalmente, enquanto dois dos informantes foram consultados através de correio eletrônico. Um dos informantes passou pelos dois tipos de consultas. As consultas por correio eletrônico foram feitas com a intenção de assegurar que a presença do pesquisador não interfere nos julgamentos dos informantes. É importante notar que os julgamentos convergiram em ambos os tipos de teste. A elicitación das sentenças como foco largo é crucial para o experimento. Isto foi testado através de um grupo de controle. Um dos informantes do grupo de controle não foi instruído sobre o tipo de leitura e seus julgamentos não convergiram com o do resto do grupo. Dois dos informantes foram consultados mais de uma vez para assegurar consistência interna. É também importante notar que este *corpus* manipula palavras bissilábicas no ambiente de retração para assegurar que, em caso de colisão acentual, retração seja a alternativa preferida para desfazer a colisão. Para outras alternativas a fim de desfazer colisão acentual, ver Abousalh (1997). Vogais tônicas altas foram evitadas neste *corpus* porque há uma aparente resistência no português de desacentuar vogais altas (e.g. caju branco -> *? caju branco).

As sentenças que compõem este *corpus* ainda não foram manipuladas acusticamente em sua totalidade. Apenas uma pequena parte dele já conta com curvas de entoação com base em extração de F0 a partir do programa Praat. No futuro próximo, todas as sentenças trabalhadas para percepção de retração de acento devem ser gravadas, e curvas de F0 devem ser extraídas para se verificar a hipótese. Para esse propósito, talvez o *corpus* seja alterado minimamente, evitando-se o uso de obstruintes surdas e controlando-se a velocidade de fala. Também espectrogramas serão feitos com base neste *corpus*, a fim de se verificar a existência de possíveis correlatos acústicos da retração acentual, como qualidade vocálica.

$$\begin{array}{ccc}
 (7) & (& x) \longrightarrow (& x) \\
 & (x) & (x)(x)(x) & (x) (x)(x)(x) \\
 & \sigma & \sigma \sigma \sigma \sigma \sigma \sigma \sigma & \sigma \sigma \sigma \sigma \sigma \sigma \sigma \\
 & \text{constantinopolitanismo} & & \text{constantinopolitanismo}
 \end{array}$$

$$\begin{array}{ccc}
 (8) & (& x) \\
 & (x) & (x)(x)(x) \\
 & \sigma & \sigma \sigma \sigma \sigma \sigma \sigma \sigma \\
 & \text{constantinopolitanismo} & &
 \end{array}$$

Em outras palavras, o modelo tradicional pode derivar apenas uma forma, que, neste caso, é uma forma inexistente (um caso de neutralização absoluta). Regras de reestruturação são necessárias para derivar variação sempre. Abaurre, Galves, Mandel & Sandalo (2001) argumentam que a teoria da otimalidade pode prever a variação livre acentual atestada sem reestruturação. O presente estudo pretende mostrar que este modelo de análise pode também evitar uma variação livre que não é atestada, prevendo a ocorrência de sintagmas fonológicos maiores e menores se a fonologia é processada paralelamente à sintaxe. Este artigo pretende mostrar que os fatos de entoação e/ou retração acentual podem ser previstos por um conflito de forças, ao invés de uma computação derivacional de sintagmas fonológicos. Entre essas forças está um princípio de euritmia que faz referência a uma relação sintática ainda não discutida nos trabalhos sobre a formação de sintagmas fonológicos em outras línguas. Segundo a hipótese de trabalho deste estudo, esta relação é a de concordância (*SPEC-HEAD agreement*).

Como os constituintes sintáticos e os constituintes prosódicos nunca são totalmente isomórficos, é de acordo comum que não há um mapeamento direto da sintaxe para a estrutura prosódica. Que tipos de informações sintáticas podem ser acessados pela prosódia, entretanto, têm sido objeto de debate. Segundo autores como Selkirk (1986), que seguem a visão do modelo de regência e vinculação (CHOMSKY, 1981), a gramática é configurada através de módulos

completamente autônomos e, portanto, a interação entre estes módulos deve ser mínima. Segundo este ponto de vista, a prosódia é processada em um componente independente, chamado de *estrutura p*, que se localiza entre a sintaxe e a fonologia, e que pode acessar informações mínimas da sintaxe, a saber fronteiras (por exemplo as fronteiras de um sintagma sintático), mas jamais relações sintáticas como c-comando, regência e concordância. Segundo Selkirk (1986), basta para a *estrutura p* um princípio como:

- (9) A fronteira {direita/esquerda} de cada $\{X^{\max}/ X^{\text{lex}}\}$ coincide com a fronteira {direita/esquerda} de um sintagma fonológico (Φ).

Há, entretanto, uma segunda linha de investigação, segundo a qual o mapeamento prosódico poderia acessar relações sintáticas. Nespor & Vogel (1986) assumem que a relação núcleo-complemento pode ser acessada. Hale & Selkirk (1987) argumentam que a regência lexical tem um papel fundamental no mapeamento tonal do tohono o'odham (anteriormente chamado de papago). Nesta língua, f e XPs são isomórficos. Mas não é qualquer XP que consiste em um domínio tonal. Os sintagmas lexicalmente governados (i.e., contidos dentro da projeção máxima de um núcleo lexical) não constituem fronteiras prosódicas. Afirmando que nem em todas as línguas é necessário fazer referência para relações sintáticas para construir um sintagma fonológico, Hale & Selkirk (1987) estabelecem o seguinte parâmetro:⁴

- (10) Alinhe a fronteira {direita/esquerda} de cada categoria sintática Scat com a categoria prosódica Pcat, onde uma língua pode escolher entre
 Scat = XP, ou
 Scat = XP somente se não for regido lexicalmente
 e Pcat = sintagma prosódico

Finalmente, há a visão da teoria da otimalidade, segundo a qual não há um mapeamento unidirecional da estrutura sintática para a estrutura prosódica, bem como não há parâmetros. Apesar de ser uma teoria gerativa, a chamada teoria da otimalidade se diferencia drasticamente da teoria chomskyana. A teoria da otimalidade, proposta inicialmente por Prince & Smolensky (1993), é uma teoria

que questiona a análise gramatical através de módulos. De acordo com este modelo, uma gramática particular é o resultado do ordenamento de um conjunto de princípios universais, isto é, aplicados em todas as línguas, mas violáveis. Como todos os princípios são processados paralelamente, pode haver um conflito entre vários deles, tornando-se impossível satisfazer a todos. Somente em uma situação de conflito um princípio pode ser violado. Uma hierarquização de princípios processados em paralelo desempenha um papel fundamental neste modelo, portanto. Truckenbrodt (1995, 1999) explora este ponto de vista e propõe que a formação de sintagmas fonológicos pode ser modelada através de uma tensão entre Embrulhe-XP (que é um princípio que requer que regência sintática seja considerada) e Alinhe-XP/ X^{lex} (que segmenta sintagmas fonológicos, acessando apenas fronteiras sintáticas à direita ou à esquerda). Assim, o autor retoma tanto a proposta de Selkirk (1986) quanto a proposta de Hale & Selkirk (1987), e as reformula em forma de princípios em conflito ao invés de parâmetros. Dependendo de onde Embrulhe-XP é ranqueado com respeito às restrições de alinhamento, diferentes segmentações prosódicas são encontradas:

(11) Tipologia segundo Truckenbrodt (1999)

Chichewa: (V NP NP)	Embrulhe-XP >> Alinhe- X^{max} à direita, Alinhe- X^{lex}
Chimwini: (V NP) (NP)	Alinhe- X^{max} à direita >> Embrulhe-XP, Alinhe- X^{lex}
Hungarian: (V) (NP) (NP)	Alinhe- X^{lex} >> Embrulhe-XP

Essa alternativa de análise acomoda as observações de Inkelas & Zec (1995) de que há línguas que parecem ser melhor explicadas pela abordagem *end-based* e outras pela abordagem *relation-based* e de que há domínios que podem ser melhor capturados pela abordagem *end-based* e outros pela abordagem *relation-based*. Se é verdade que em uma mesma língua o mapeamento de diferentes domínios segue condições distintas, em uma abordagem paramétrica seremos forçados a acreditar que uma mesma língua pode mudar o parâmetro estabelecido para cada um dos domínios a ser processado, o que esvazia uma visão paramétrica de diversidade lingüística. Em uma teoria representacional, é possível hierarquizar princípios que

acessem relações sintáticas relevantes para outros domínios mais acima ou mais abaixo que o princípio de alinhamento, gerando o fato de certos domínios levarem em conta apenas fronteiras sintáticas e outros levarem em conta relações sintáticas.

O PB apresenta um fenômeno que envolve acesso a uma relação sintática não prevista pela tipologia de Truckenbrodt (1999). Esse fenômeno foi notado pela primeira vez por Sandalo (2000) com base em entoação e discutido por Sandalo & Truckenbrodt (2002) com base em entoação e retração de acento. O texto de maior destaque na literatura internacional relatando a presença de retração acentual no PB foi o de Major (1985). O foco do texto de Major, entretanto, não é o fenômeno de retração de acento e nem mesmo o estudo sobre fronteiras prosódicas. Major (1985: 280) meramente afirma que “there is a slight tendency for stress to shift to alternating patterns in a sentence in Brazilian Portuguese”. Major oferece um único exemplo em que o acento pode ou não retrair, e o autor desconsidera totalmente casos em que o acento não pode retrair, os quais são fundamentais para se estabelecer fronteiras prosódicas no quadro gerativo. Assim, não importa para o presente estudo se uma retração efetivamente ocorre ou não (isto é, que ela não seja obrigatória). O dado crucial para se estabelecer fronteiras prosódicas é aquele no qual uma retração em situação de colisão acentual é percebido como inaceitável, isto é, onde a retração é bloqueada. Como mostrado em 12, o acento de uma palavra como *café* pode retrair para a sílaba anterior quando esta palavra está diante de outra palavra cuja primeira sílaba é acentuada (as sílabas sublinhadas foram percebidas como acentuadas pelos nossos informantes). Mas os exemplos em 13 e 14 mostram que nem sempre esta retração é permitida. Os símbolos * e *? representam realizações rejeitadas pelos falantes nativos consultados, em que * representa uma realização ainda pior que aquela marcada por *?.

(12) Café quente queima a boca -> Café quente queima a boca

(13) Café quente queima -> *? Café quente queima

(14) O novo café queima -> * O novo café queima

É largamente assumido pela fonologia prosódica que retração acentual pode ocorrer dentro de um domínio prosódico, mas nunca através de fronteiras de domínios prosódicos (cf. NESPOR & VOGEL, 1986). Assim, devemos entender que há uma fronteira fonológica entre *café* e seu complemento em (13), em que a estrutura é N A V. Em (12), entretanto, onde a estrutura sintática é N A V N, o sintagma nominal *café quente* forma um único sintagma fonológico, uma vez que a retração é permitida. Isto é, não temos variação livre e temos os mesmos domínios notados para a entoação mencionados na seção 1 deste projeto.

Dados apenas os exemplos 12-14 acima, alguém poderia argumentar que o PB tem uma forte preferência por binariedade na formação de sintagmas fonológicos. Assim, seqüências de duas palavras prosódicas seriam agrupadas em um mesmo sintagma fonológico sempre que a sentença como um todo tenha um número par de palavras. Esta hipótese, entretanto, prevê que uma sentença como *Café queima* seria mapeada em um único sintagma fonológico e não é o que ocorre. Uma fronteira prosódica é sempre encontrada no português brasileiro entre o sujeito e o predicado, uma vez que a retração é sempre bloqueada entre o sujeito e o predicado, como pode ser notado em 14 e também nos exemplos abaixo. Um algoritmo com base em binariedade apenas não pode prever este fato. Em todos os exemplos no restante deste texto, os parênteses, “()”, representam sintagmas fonológicos. Os sintagmas sintáticos são representados por XPs, onde X é uma variável.

(15) [N]_{NP} V

(N)(V)

Café queima.

* Café queima

(N)(V)

O José dança.

* O José dança

(16) [N]_{NP} V [N]_{NP}

(N) (V
 Café queima a boca.
 * Café queima

(N) (V
 José come uvas.
 * José come

(17) [N A]_{NP} V

N A) (V
 O cavalo chinês corre.
 * chinês corre

A N) (V
 O novo café queima.
 * café queima

(18) [DET N A]_{NP} V [N]_{NP}

N A) (V
 O cavalo chinês come uvas.
 * chinês come

A N) (V
 O novo café queima a boca.
 * café queima

Os dados acima indicam que Alinhe-XP (SELKIRK, 1986) é importante para a análise, porque uma fronteira prosódica é sempre estabelecida à direita da fronteira do sintagma nominal (NP) sujeito. Evidência de que Alinhe-XP é importante para a análise do PB vem também de dados que contam com sintagmas adverbiais (ADVP). Em 19 há um sintagma adverbial separando o sujeito e o verbo e em 20 há um sintagma adverbial entre o verbo e o objeto direto. Em ambos os casos, uma fronteira prosódica demarca a fronteira direita do sintagma adverbial (uma vez que retração é bloqueada), como previsto pelo princípio Alinhe-XP.

- (19) [DET N A]_{NP} [Adv]_{AdvP} V [N]_{NP}
 Adv) (V
 A abelha rainha amanhã compra livros.
 * amanhã compra

- (20) [DET N A]_{NP} V [Adv]_{AdvP} [N]_{NP}
 Adv) (N
 A abelha rainha comerá amanhã uvas.
 * amanhã uvas

O fato de um advérbio poder aparecer entre o verbo e seu objeto direto sugere que o verbo se moveu de sua posição inicial para a posição nuclear de uma projeção funcional mais acima. O sintagma adverbial poderia, então, estar adjungido à esquerda do VP, seguindo linearmente o verbo após movimento (COSTA, 1998: 19-36). Este cenário não afeta os casos derivados até agora porque a fronteira direita do verbo não é de natureza frasal e, portanto, não invoca Alinhe-XP.

O princípio de alinhamento *Alinhe XP* à direita não dá conta do contraste em 12 e 13, entretanto. Também Embrulhe-XP não tem nada a dizer sobre estes fatos, porque este princípio, se está alto em uma dada hierarquia, jamais permitiria a separação de um núcleo e seu complemento como ocorre em 13. Este exemplo não é um caso isolado do PB, como pode ser notado no contraste entre 21 e 22. Estes exemplos comparam a seqüência sujeitos N A em sentenças transitivas e intransitivas e notam que retrações aceitáveis quando o verbo é transitivo passam a ser inaceitáveis quando intransitivo:

- (21) (N_A) (V N)
Café quente esquenta o corpo.
 Um chinês louco dançou samba.

- (22) (N A) (V) -> (N)(A) (V)
 (N) (A) (V)
Café quente esquenta.
 *? Café quente.
 O Café quente sumiu.
 *O Café quente
 Um chinês louco morreu.
 * Um chinês louco

Não só a transitividade do verbo tem um papel importante, mas também a do sujeito. Um sujeito sem complemento força a criação de uma fronteira prosódica entre um verbo transitivo e seu argumento interno, como pode ser notado nos pares mínimos em 23 e 24:

(23) (N A) (V_N)
 O canguru australiano dançou samba.

(N A) (V_N)
 A abelha rainha comeu uvas.

(24) (N) (V N) -> (N) (V) (N)

(N) (V)(N)
 O canguru dançou samba.
 * dançou samba

(N)(V)(N)
 Pedro comeu uvas.
 * comeu uvas

Os dados indicam que o sintagma fonológico carregando o verbo deve estar refletido no sujeito e vice-versa, como esquematizado em (25), que representa o mapeamento prosódico sugerido pelos dados de percepção de retração acentual de Sandalo & Truckenbrodt (2002). Mais exemplos seguem em 26 a 28.

(25) (N A)(V N)(Adv)
 (N A)(V Adv)(N)

(N A)(V N)(Adv)
 (N A)(V Adv)(N)
 (N A)(V O1)(O2)
 (N)(V)(N A)

(26) [N A]_{NP} V [Adv]_{AdvP} [N]_{NP}

a. (N_A) (V Adv) (N)
 Um café quente queimou ontem a boca.

(N_N) (V Adv) (N)
José Carlos come sempre maçãs.

Mané Pedro compra sempre livros.

- b. (N A)(V_Adv)(N)
A abelha rainha comeu ontem maçãs.
- (27) [N X]_{NP} V [N1]_{NP} [N2]_{NP}
- a. (N_N)(V N1)(N2)
José Carlos deu um café p(a)ra Maria.
- b. (N N)(V_N1) (N2)
José Carlos comprou tudo p(a)r(a) o filho.
- c. (N N)(V_N1) (N2)
Ana Maria pegou água da bica.
- (28) [N]_{NP} V [N X]_{NP}
- a. (N)(V)(N A)
José bebe café quente.
* José bebe
- b. (N)(V)(N A)
A abelha comeu uvas brancas.
* comeu uvas
- c. (N) (V) (N_A)
O canguru bebe café quente.
- (N)(V)(N_N)
O homem viu José Carlos.

Sandalo & Truckenbrodt propõem que os fatos do português brasileiro podem ser derivados pela interação entre Alinhe-XP e um princípio de euritmia que rotulamos de uniformidade, que é o princípio não dominado.

(29) Uniformidade (Sandalo & Truckenbrodt (2002):

Sujeito e verbo (se adjacentes) são mapeados em unidades do mesmo tamanho.

Alguém poderia ainda tentar argumentar que não se trata do princípio em 29, mas que o PB força os dois primeiros sintagmas fonológicos a terem o mesmo tamanho/peso, sendo um princípio

puramente fonológico que não tem nenhuma relação com a sintaxe. Os dados abaixo sugerem que esta análise não está correta:

- (30) $[[N X]_{NP} \& [N X]_{NP}]_{NP} V$
 (N A) & (N N)(V) -> (N)(A) & (N)(N) (V)
 (N) (A) & (N)(N)(V)
 O café quente e a Ana Raquel sumiram.
 *? café quente
 O José Carlos e a Ana Maria morreram.
 *? José Carlos

Os dados em 30 são importantes para se manter 29, porque indicam que o sintagma fonológico carregando o verbo, por conter apenas uma palavra lexical, força o sujeito complexo e coordenado a ser mapeado em quatro sintagmas fonológicos que contêm apenas uma palavra lexical cada. Não se trata de garantir o mesmo tamanho/peso para os dois primeiros sintagmas fonológicos. Se este fosse o caso, cada sujeito complexo seria mapeado em um mesmo sintagma fonológico. Os dados em 30 indicam que o sujeito e o verbo necessitam ser acessados.

Note que este estudo tem uma importante contribuição teórica. Se o princípio em 29 se mostrar verdadeiramente ativo no PB a partir de um estudo no qual mais situações sintáticas são avaliadas (considerando retração de acento e entoação), teremos um argumento para acreditar que a fonologia não é meramente interpretativa e deve ser processada paralelamente à sintaxe. Isto porque o princípio 29 faz referência a uma relação sintática, a saber, SUJEITO. Isto sugere que o mapeamento sintaxe-fonologia necessita de mais informação do que se tem assumido. Este mapeamento requer informação não apenas sobre rótulos de categorias e fronteiras de sintagmas sintáticos, como se é tradicional aceitar, mas também sobre conceitos relacionais como SUJEITO. Além disso, estes dados colocam uma questão interessante: por que SUJEITO? Sujeito não é uma noção primitiva na teoria gerativa; é derivado estruturalmente. Os dados abaixo sugerem uma resposta para esta pergunta, que este projeto tem como objetivo investigar mais amplamente (nos dados abaixo V representa verbo, O representa objeto e S representa sujeito).

- (31) **VtransitivoO**
Comprou tudo
Ganhou todos
- (32) **VinacusativoS**
 * Rolou tudo
 * Chegou tudo
- (33) **VinergativoS**
 * Nadou tudo
 * Correu Pedro
- (34) **VinergativoSS**
 Correu o José Carlos
 Nadou Café Quente
- (35) **SSVinergativo**
 * José Carlos correu
 * Café Quente nada
 * José Carlos nada
- (36) **SSVinacusativo**
 * Um café quente rolou
 * O sofá branco quebrou
 * O José Carlos chegou
- (37) **VinacusativoSS**
 Chegou o José Carlos
 Quebrou o sofá branco
 Rolou um café quente

Note que os dados 31-33 reforçam a conclusão de que SUJEITO é aqui uma noção importante, uma vez que o PB se comporta de maneira diferente caso o sintagma nominal que segue o verbo seja um objeto ou um sujeito posposto. Os dados acima têm algo mais a sugerir. Os dados 34-37 mostram que o sujeito se comporta de maneira distinta se ele está em sua posição canônica antes do verbo ou se está posposto. Se o sujeito está em sua posição canônica, uniformidade se aplica (35

e 36). Se o sujeito é posposto, uniformidade não se aplica e o mapeamento prosódico é dado por Alinhe-XP, como pode ser notado em 34 e 37. Esse conjunto de dados corrobora a análise sintática de Kato & Tarallo (1993) e Kato & Raposo (1994), uma vez que esses autores argumentam que um sujeito posposto está deslocado (isto é, não está adjacente ao verbo) e que a posição regular do sujeito (o especificador do IP) é preenchida por *pro*. A análise de Kato & Tarallo e Kato & Raposo é sustentada pelo fato de sujeitos pospostos não obrigarem a uma concordância verbal de terceira pessoa do plural no português coloquial (cf. *chegou estas cartas hoje*). O sujeito parece concordar com uma terceira pessoa singular (*pro*). Segundo a formulação de uniformidade de Sandalo & Truckenbrodt (29), adjacência sujeito-verbo é fundamental para a aplicação de uniformidade. A hipótese que este estudo deve explorar é a de que adjacência sujeito-verbo é importante porque a fonologia está acessando o domínio da concordância do português (*SPEC-HEAD agreement*). Se esta hipótese se provar adequada, teremos um ponto bastante forte para questionar a visão tradicional de fonologia, segundo a qual algoritmos prosódicos acessam apenas informações sintáticas bastante limitadas. O PB parece apresentar um fenômeno fonológico que acessa uma das principais relações sintáticas: concordância. Em trabalhos futuros, pretendo analisar sentenças envolvendo verbos no gerúndio. O gerúndio não envolve concordância (e.g. *Um café quente queimando a boca é indesejável*) e, assim, a previsão é de que uniformidade deixe de ativar.

Além de sua (potencial) contribuição teórica, este estudo já conta com um avanço na descrição de prosódia. Em outras palavras, um aspecto positivo do estudo já realizado é poder gerar frases fonológicas do português brasileiro sem regras de reajuste. A eliminação de regras de reajuste significa uma simplificação para a teoria fonológica. Resta testar mais amplamente o modelo proposto e apresentar mais evidências para uma fonologia menos interpretativa (isto é, que pode acessar vários tipos de relações sintáticas). Finalmente, resta analisar a prosódia de outras línguas e investir mais na busca de uma configuração de gramática capaz de prever os fenômenos de interface existentes e excluir outros tipos de interface sintaxe-fonologia que jamais ocorrem.

NOTAS

¹ Este artigo resume duas participações no IV Congresso Nacional de Fonética e Fonologia/I Congresso Internacional de Fonética e Fonologia: (i) Algumas Observações sobre a Formação de Sintagmas Fonológicos em Português Brasileiro: um estudo com base em entoação e desfazimento de colisão acentual, e (ii) Retração de Acento e a Posição do Sujeito Gramatical no Português Brasileiro.

² Os falantes consultados são de São Paulo, Minas Gerais e Espírito Santo.

³ Um estudo detalhado, entretanto, sobre entoação a partir da extração de F0 (isto é, a partir de dados acústicos) é crucial para este estudo. Como observado por Eleonora Albano (comunicação pessoal), é possível que não haja verdadeira retração acentual no PB; o que é percebido como retração pode ser uma impressão auditiva causada pela mudança da curva entoacional. Este estudo buscará, portanto, no futuro, verificar se existem correlatos acústicos para a retração acentual, seguindo a linha de trabalho de Madureira & Barbosa (2001). Isto é, um dos objetivos da análise acústica será verificar se retração acentual e entoação são ou não fenômenos independentes.

⁴ O parâmetro estabelecido por Hale & Selkirk foi extraído de Trukenbrodt (1999), que simplifica a redação original.

REFERÊNCIAS

ABAURRE, M. B.; GALVES, C.; MANDEL, A.; SANDALO, F. 2001. The Sotaq optimality based computer program and secondary stress in two varieties of Portuguese. Rutgers Optimality Archive <<http://roa.rutgers.edu>>.

ABOUSALH, E. S. F. *Resolução de choques de acento no português brasileiro*. 1997. Dissertação (Mestrado) - Universidade Estadual de Campinas.

ATKINSON, M. *Children's Syntax*. Cambridge, MA: Blackwell, 1992.

BISOL, L. *O acento: duas alternativas de análise*. 1992. (Ms.) - Pontifícia Universidade Católica do Rio Grande do Sul.

CHEN, M. Y. The syntax of Xiamen tone sandhi. *Phonology Yearbook*, n. 4, p. 109-149, 1987.

CHOMSKY, N. *Lectures on Governing and Binding*. Dordrecht: Foris, 1981.

CHOMSKY, N. *The minimalist program*. Cambridge, Mass.: MIT Press, 1995.

COLLISCHONN, G. *Um estudo do acento secundário em português*. 1993. Dissertação (Mestrado) - Universidade Federal do Rio Grande do Sul.

- COSTA, J. *Word order variation*. 1998. Tese (Doutorado) – Holland Institute of Generative Linguistics.
- DE LACY, P. *Higher prosody in Maori*. 1999. Ms., University of Massachusetts, Amherst.
- GALVES, C. O objeto nulo no português brasileiro: percurso de uma pesquisa. *Cadernos de Estudos Lingüísticos* 17, 1989.
- FROTA, S. *Prosody and focus in European Portuguese*. 1998. Tese (Doutorado) – Universidade de Lisboa.
- FROTA, S.; VIGÁRIO, M. Aspectos de prosódia comparada: Ritmo e entoação no pe e no pb. 2000. Ms., Universidade de Lisboa, and Universidade do Minho, Braga, Portugal.
- GERKEN, LouAnn. Prosodic structure in young children's language production. *Language* 72, 1996.
- GHINI, M. Ø-formation in Italian: A new proposal. In: DYCK, C. (Ed.). *Toronto Working Papers in Linguistics*, v. 12, n. 2. Department of Linguistics, University of Toronto, p. 41-78, 1993.
- GUSSENHOVEN, C. The English rhythm rule as an accent deletion rule. *Phonology* 8, v. 1, p. 1-35, 1991.
- HALLE, M., VERGNAUD, J.-R. *An essay on stress*. 1987. Cambridge, Mass.: MIT Press.
- HARRIS, J. W. The Stress Erasure Convention and Cliticization in Spanish. *Linguistic Inquiry* 20, p. 339-363, 1989.
- HAYES, B. The phonology of rhythm in English. *Linguistic Inquiry* 15, v. 1, p. 33-74, 1984.
- HAYES, B. The prosodic hierarchy in meter. In: KIPARSKY, P.; YOUMANS, G. (Ed.) *Rhythm and meter*, Orlando: Academic Press, 1989.
- HAYES, B. *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press, 1995.
- HAYES, B.; PUPPLE, S. On the rhythm rule in Polish. In: HULST, H. van der; SMITH, N. (Ed.). *Advances in nonlinear phonology*, Dordrecht: Foris, p. 59-81, 1985.
- HAYES, B.; LAHIRI, A. Bengali intonational phonology. *Natural Language & Linguistic Theory* 9, v. 1, p. 47-96, 1991.
- HSIAO, F. *Tonal domains are stress domains in Taiwanese—evidence from focus*. 2001. Ms., Massachusetts Institute of Technology.

INKELAS, S.; ZEC, D. Syntax-phonology interface. In: GOLDSMITH, J. (Ed.). *The handbook of phonological theory*. Cambridge, MA: Blackwell, 1995. p. 535-549.

JACKENDOFF, R. S. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press, 1972.

KANERVA, J. M. *Focus and phrasing in Chichewa phonology*. (1989). Tese (Doutorado) – Stanford University.

KANERVA, J. M. Focus on phonological phrases in Chichewa. In: INKELAS, S.; ZEC, D. (Ed.). *The phonology-syntax connection*. Chicago: The University of Chicago Press, 1990. p. 145-161.

KATO, M.; TARALLO, F. The loss of VS syntax in Brazilian Portuguese. In: KOCH, I. V.; SCHLIEBEN-LANGE, B. (Ed.). *Linguistik in Brasilien*. Tübingen, Niemeyer, 1993.

KATO, M.; RAPOSO, E. European and Brazilian word order; questions, focus and topic constructions. In: PARODI, C.; QUICOLI, A.C.; ZUBIZARRETA, M.L. (Ed.). *Romance Linguistics in Los Angeles*. Georgetown University Press, 1994.

LADD, D. R. *Intonational Phonology*. Cambridge: CUP, 1996.

LEBEAUX, D. *Language acquisition and the form of the grammar*. 1988. Tese (Doutorado) – University of Massachusetts.

LÉVI-STRAUSS, C. *Tristes Trópicos*. São Paulo: Companhia das Letras, 1955/2000.

LIBERMAN, M.; PRINCE, A. On stress and linguistic rhythm. *Linguistic Inquiry* 8, v. 2, p. 249-336, 1977.

MAJOR, R. C. Stress-timing in Brazilian Portuguese. *Journal of Phonetics* 9, p. 343-351, 1985.

MADUREIRA, S.; BARBOSA, P. Seminário do LAFAPE, out. 2001

McCARTHY, J.; PRINCE, A. Generalized alignment. In: BOOIJ, G.; MARLE, J. van (Ed.). *Yearbook of morphology 1993*, Dordrecht: Kluwer, 1993. p. 79-153.

MÉTRAUX, A. Ethnography of the Chaco. In: STEWARD, J. H. (Ed.). *Handbook of South American Indians*, v. 1: p. 197-310, 1945. Washington: Government Printing Office (Bulletin 143, Bureau of American Ethnology, Smithsonian Institution)

NESPOR, M. On the separation of prosodic and rhythmic phonology. In: INKELAS, S.; ZEC, D. (Ed.). *The phonology-syntax connection*. Chicago: The University of Chicago Press, 1990. p. 243-258.

NESPOR, M. Stress domains. In: HULST, H. van der (Ed.). *Word prosodic systems in the languages of Europe*. Berlin: Mouton de Gruyter, 1999. p. 117-159.

- NESPOR, M.; VOGEL, I. Clash avoidance in Italian. *Linguistic Inquiry* 10, v. 3, p. 467-482, 1979.
- NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris, 1986.
- NESPOR, M.; VOGEL, I. On clashes and lapses. *Phonology* 6, v. 1, p. 69-116, 1989.
- PRINCE, A. Relating to the Grid. *Linguistic Inquiry* 14, p. 19-100, 1983.
- PRINCE, A.; SMOLENSKY, P. *Optimality Theory: Constraint interaction in generative grammar*. 1993. Ms., Rutgers University and University of Colorado, Boulder.
- ROCA, I. M. Secondary Stress and Metrical Rhythm. *Phonology Yearbook* 3, p. 341-370, 1986.
- RODRIGUES, A. *Línguas Brasileiras*. São Paulo: Loyola, 1986.
- SANDALO, F. *A Grammar of Kadiwéu with special emphasis to the polysynthesis parameter*. 1997. MIT Occasional Papers in Linguistics 11.
- SANDALO, F. A prosódia da escrita em textos contemporâneos e clássicos. Simpósio Considerações sobre a Escrita Alfabética: Relação entre Aquisição do Sistema e Conhecimento do Componente Fônico da Língua. V Encontro Nacional sobre Aquisição da Linguagem e I Encontro Internacional sobre Aquisição de Linguagem, Porto Alegre, 2-6 out. 2000.
- SANDALO, F.; TRUCKENBRODT, H. Some Notes on Phonological Phrasing in Brazilian Portuguese. In: CSIRMAZ, A.; LI, Z.; NEVINS, A.; VAYSMAN, O.; WAGNER, M. *Phonological Answers: MIT Working Papers in Linguistics* 42, p. 285-310, 2002.
- SANTOS, R. S. *A Aquisição do Acento Primário no Português Brasileiro*. 2001. Tese (Doutorado) - Universidade Estadual de Campinas.
- SCARPA, E. Sons preenchedores e guardadores de lugar: relações entre fatos sintáticos e prosódicos na aquisição da linguagem. *Estudos de Prosódia*. Campinas: Editora da UNICAMP, 1999.
- SELKIRK, E. On derived domains in sentence phonology. *Phonology Yearbook* 3, p. 371-405, 1986.
- SELKIRK, E. The prosodic structure of function words. In: BECKMAN, J.; DICKEY, L. W.; URBANCZYK, S. (Ed.). *Papers in Optimality Theory. University of Massachusetts Occasional Papers* 18, Amherst, Mass: GLSA, p. 439-469, 1995.
- SELKIRK, E.; SHEN, T. Prosodic domains in Shanghai Chinese. In: INKELAS, S.; ZEC, D. (Ed.). *The phonology-syntax connection*. Chicago: The University of Chicago Press, 1990. p. 313-337.

SELKIRK, E.; TATEISHI, K. Syntax and downstep in Japanese. In: GEORGIOPOULOS, C.; ISHIHARA, R. (Ed.). *Interdisciplinary approaches to language: Essays in honor of s -y Kuroda*. Dordrecht: Kluwer, 1991. p. 519-543.

SHATTUCK-HUFNAGEL, S.; OSTENDORF, M.; ROSS, K. Stress shift and early pitch accent placement in lexical items in American English. *Journal of Phonetics* 22, p. 357-388, 1994.

TRUCKENBRODT, H. *Phonological phrases: Their relation to syntax, focus, and prominence*. 1995. Tese (Doutorado) – Massachusetts Institute of Technology.

TRUCKENBRODT, H. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30, v. 2, p. 219-255, 1999.

VIGÁRIO, M. *The Prosodic Word in European Portuguese*. 2001. Tese (Doutorado) – Universidade de Lisboa.

VISCH, E. *A metrical theory of rhythmic stress phenomena*. 1989. Tese (Doutorado) – University of Utrecht.

Retração acentual e verbos transitivos com leitura intransitiva

Raquel S. Santos
Universidade de São Paulo

Abstract

The focus of this paper is on the relation between syntax and prosody, by discussing the role of the empty syntactic categories in the application of the stress shift rule. The discussion about the influence of empty categories on phonological processes is not new. Usually, phonology papers argue that the access to syntax is indirect, and there is no access to phonological empty syntactic categories (cf. NESPOR; VOGEL, 1986). However, Santos (2002) argues that phonological empty syntactic categories can intervene on the rhythmic organization of utterances. She defends that *pro* intervenes on stress shift rule, while trace is not taken into account for that rule. In order to show that, she tests utterances with stress clash between verb and adverb.

In this presentation, the syntactic structures with verbs which can have a transitive or intransitive reading are tested with *pro*. The hypothesis is that when there is a transitive reading, the stress clash can not be solved. In the sentences where the intransitive reading is possible, there is no *pro* between the verb and adverb, and the stress shift rule is allowed. This prediction can be used as an evidence which favors Galves (1989) analysis that null object in Brazilian Portuguese is a *pro*, and not a trace.

1. INTRODUÇÃO

A discussão sobre a influência que os diversos componentes da gramática (fonologia, morfologia, sintaxe, semântica) podem exercer entre si não é nova nos estudos lingüísticos e nem deixou de ser observada nas diferentes correntes teóricas. A grande questão é definir, para cada língua, quais elementos de um determinado componente X influenciam um componente Y. Uma vez que a influência ocorre, é plausível assumir que haja indícios da mesma; isto é, encontrar, em Y, indícios de que X é levado em conta.

A fonologia sempre foi um lugar onde a influência de outros componentes gramaticais é claramente apontada. Regras fonológicas segmentais fazem referência a fronteiras de morfemas ou de palavras. Regras como marcação de foco levam em conta informações semânticas e sintáticas. A acentuação depende da classe de palavras.

As explicações para a relação entre os diversos componentes gramaticais têm sido várias, e mais variadas ainda para os fenômenos ditos prosódicos. Chomsky & Halle (1968) propõem que as regras fonológicas dependem dos tipos de fronteiras entre as palavras; Jaeggli (1980) argumenta que o traço formal de caso é considerado para a aplicação da regra de *wanna-contraction*, Kaisse (1985) defende que as regras fonológicas dependem exclusivamente de relações sintáticas – especificamente a relação de m-comando – entre os itens lexicais; Selkirk (1984) e Nespor & Vogel (1986) argumentam que as regras fonológicas ocorrem em domínios prosódicos que são formados pela interação entre os diversos componentes da gramática.

Discutindo a resolução de choques acentuais, Frota (1995) e Abousalh (1997) defendem que esta resolução necessita da noção de domínios prosódicos, mas que também é sensível a estruturas sintáticas. Santos (2002) argumenta que a regra de retração acentual leva em conta estruturas sintáticas fonologicamente vazias (*pro*).

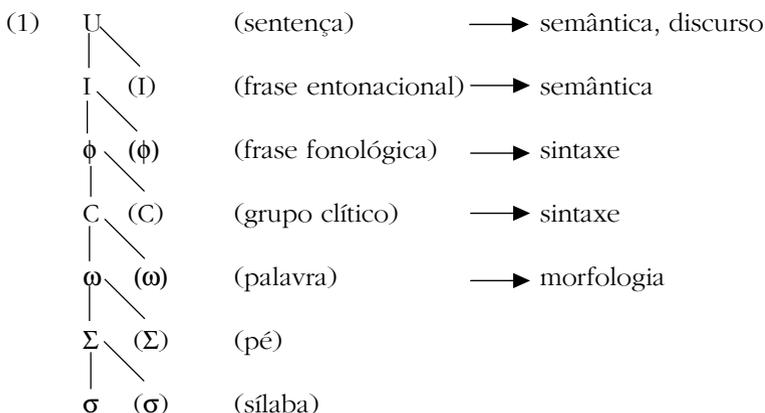
Neste artigo, analiso como se comportam, em relação à retração acentual, as estruturas de sentenças simples formadas por verbos intransitivos, transitivos com leitura intransitiva, transitivos com objeto realizado, transitivos com objetos nulos e verbos que formam um predicado complexo. Além dos resultados sobre o comportamento fonológico dessas estruturas, esta análise traz evidências a favor da análise de Galves (1989) de que o objeto nulo, em português brasileiro, é um *pro*.

Este artigo está organizado da seguinte maneira: na seção 3 é discutida a proposta de Nespor & Vogel para o mapeamento sintático e para retração acentual. A 4ª. seção traz as propostas sintáticas para as categorias sintáticas vazias para o português. A 5ª seção trata dos cuidados metodológicos deste trabalho. Na seção 6 analisamos os dados, envolvendo categorias sintáticas vazias e aplicação de regras rítmicas. Finalmente, na seção 7 resumimos nossas observações.

2. PRESSUPOSTOS TEÓRICOS

Segundo Nespor & Vogel (1986), as representações mentais da fala são organizadas hierarquicamente em constituintes prosódicos, que servem como domínio de aplicação de processos fonéticos e regras fonológicas. Alguns desses domínios são construídos levando em conta outros componentes gramaticais. No que diz respeito especificamente à sintaxe, Nespor & Vogel afirmam haver duas hipóteses que consideram o acesso direto da fonologia. Na versão forte, todos os elementos foneticamente vazios (vestígio de NP, PRO, vestígio-*Wh*, etc) teriam o mesmo status que itens lexicais com realização fonética e bloqueariam adjacência; já na versão mais fraca, apenas os elementos marcados por Caso seriam capazes de bloquear certas operações fonológicas. As autoras defendem que todos os casos discutidos na literatura podem ser analisados apenas com base na hierarquia prosódica, sem acesso às informações sintáticas e, por isso, defendem um acesso indireto aos outros componentes gramaticais; isto é, para a construção de um determinado domínio

prosódico, leva-se em conta a informação de outros componentes. Uma vez construído esse domínio, não há mais acesso às informações dos outros componentes gramaticais. As autoras propõem sete níveis prosódicos, representados em (1) com os respectivos componentes gramaticais com que se relacionam:



As operações fonológicas devem levar em conta o nível em que ocorrem. Outra distinção necessária é o local em que os processos fonológicos ocorrem. Há processos fonológicos que ocorrem em junturas de domínios, outros nas fronteiras do domínio, e por fim os que ocorrem no interior do domínio (SELKIRK, 1980, NESPOR & VOGEL, 1986).

Frota (1995) e Abousalh (1997) argumentam que a retração acentual ocorre em f no português, tanto europeu quanto brasileiro, o que significa que leva em conta informações sintáticas. Segundo as autoras, trata-se de uma regra opcional, e muitas outras estratégias podem ser utilizadas com o mesmo intuito de eliminar a colisão acentual (pausa, alongamento, alternância de tons). As autoras também propõem que a retração acentual é uma regra que ocorre dentro dos domínios desse nível (e não nas fronteiras dos mesmos).

Nespor & Vogel propõem o seguinte argumento para a construção dos domínios em ϕ :

(2) ϕ

- I. ϕ domain: The domain of ϕ consists of a C which contains a lexical head (X) and all Cs on its nonrecursive side up to the C that contains another head outside the maximal projection of X.
- II. ϕ construction: Join into an n-ary branching ϕ all Cs included in a string delimited by the definition of the domain of ϕ .
- III. 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.

De acordo com este algoritmo, uma sentença como “o José Maria comeu bolo” deve ter os seguintes domínios:¹

(3)
$$\begin{array}{l} (x)(x)(x)(x)\phi \longrightarrow (x)(x)\phi \\ (x)(x)(x)(x)C \quad (x)(x)(x)(x)C \\ \text{[}_{\text{p}}\text{ o José Maria [}_{\text{r}}\text{ [}_{\text{l}}\text{ comeu] bolo]]} \quad \text{o José Maria comeu bolo} \end{array}$$

Santos (2002), no entanto, defende que, mesmo depois de formados os domínios, a fonologia tem acesso a informações sintáticas. Os argumentos da autora são baseados no fato de que as duas sentenças abaixo (4) e (5) devem ter a mesma configuração dos domínios, dado o algoritmo de Nespor & Vogel e, no entanto, só a sentença (4) permite a retração acentual:²

$$\begin{array}{l} (x)(x)(x)(x)\phi \longrightarrow (x)(x)\phi \\ (x)(x)(x)(x)C \quad (x)(x)(x)(x)C \end{array}$$

(4) o José Maria canTOU HOje \longrightarrow o José Maria CANtou HOje

(5) o José Maria conTOU HOje \longrightarrow # o José Maria CONtou HOje

Segundo Santos, a impossibilidade de retração acentual em (5) deve-se ao fato de que, entre o verbo e o advérbio, ocorre um *pro*. Assim, o que parece ser uma adjacência acentual na realidade não o é, pois há um *pro* entre o verbo e o advérbio na sentença (5).

No entanto, Santos (2002) não analisou sentenças simples com objetos nulos. Sua análise baseou-se em sentenças com ilhas e sentenças simples com foco ou tópico, de forma a garantir que o objeto nulo não seria movido para uma posição mais alta na estrutura sintática.

Sentenças com ilha impedem o movimento do objeto para fora das mesmas. Uma sentença com ilha e foco é gramaticalmente ruim, o que indica que o foco é resultado de movimento (6a) e há um vestígio preenchendo a posição de objeto (6b). Por outro lado, sentenças com ilha e tópico são boas, o que indica que a estrutura de tópico foi gerada *in situ* (7a) e que há um *pro* preenchendo a posição de objeto (7b):

(6a) * [só esse livro]_{FOCO} eu tive a confirmação [de que o Carlos perdeu __ hoje]

(6b) * [só esse livro]_{FOCO} eu tive a confirmação [de que o Carlos perdeu *t* hoje]

(7a) [esse cargo]_{TOP} eu tive a confirmação [de que o Carlos perdeu __ ontem]

(7b) [esse cargo]_{TOP} eu tive a confirmação [de que o Carlos perdeu *pro* ontem]

A sentença (7b) pode ser testada quanto à retração acentual, pois é uma sentença gramaticalmente bem formada. Esse tipo de sentença não permite a retração acentual, indicando que *pro* é considerado pela regra de resolução de encontro (8):

(8) # esse cargo eu tive a confirmação [de que o Carlos PERdeu *pro* ONtem]

As sentenças com ilha e foco não podem ser testadas fonologicamente porque já são gramaticalmente ruins. No entanto, sabendo que foco envolve movimento, é possível testar sentenças simples com foco e observar que a retração acentual pode ocorrer (9):

(9) [só esse livro]_i o João disse que a Maria COMprou *t_i* ONtem.

3. PROPOSTAS SINTÁTICAS PARA O OBJETO NULO EM PORTUGUÊS

As propostas de análise para o objeto nulo em português não são concordes. Raposo (1986), baseado na inaceitabilidade de sentenças contendo objeto nulo dentro de ilhas (10 e 11), propôs que este é um vestígio, resultado de movimento. Em Raposo (1998), o autor modifica sua proposta e assume que, tanto para o português europeu quanto para o português brasileiro, o objeto nulo é uma categoria mista, um DP com D nulo e complemento *pro*, que se

move para foco. As duas propostas assumem movimento, o que significa que uma sentença como (5) deve ser analisada como em (12):

(10) * O rapaz que trouxe __ mesmo agora da padaria era o teu afilhado.

(11) * Que a IBM venda __ a particulares surpreende-me.

(12) OP_i o José maria conTOU t_i HOje.

As análises sobre o português brasileiro, no entanto, mostram que muitas sentenças consideradas agramaticais para o português europeu são aceitáveis para o português brasileiro (cf. GALVES, 1989, KATO, 1993, CYRINO, 1994, FERREIRA, 2000, RAPOSO & KATO, 2001). Nessa língua, as sentenças (10) e (11) são plenamente aceitáveis. Galves (1989) propõe que o objeto nulo em português brasileiro é um *pro* gerado na base e ligado a um sujeito externo, não precisando de um operador nulo para ser licenciado.³ De acordo com essa análise, a representação da sentença (5) em PB deve ser como em (13), e não como em (12):

(13) o José Maria conTOU *pro* HOje.

Essencialmente, as duas propostas divergem quanto a considerar o objeto nulo como um vestígio ou como *pro*. Uma vez que Santos (2002) mostrou que essas categorias vazias comportam-se diferentemente quando da aplicação da regra de resolução de encontro acentual, a possibilidade ou não de aplicação dessa regra constitui evidência externa para uma ou outra análise do objeto nulo em português.

4. METODOLOGIA

Para este trabalho, foi constituído um *corpus* de 28 sentenças em que se testou a possibilidade de retração acentual com 4 falantes do português brasileiro, resultando em 112 sentenças. Os testes foram perceptuais.⁴ As sentenças eram lidas pelo pesquisador e os informantes deveriam dizer se a retração era possível ou não.

Uma vez que o interesse era observar se os tipos de verbos influenciariam a possibilidade de retração acentual, cada sentença

do *corpus* era a resposta para uma pergunta que se inseria numa situação descrita. Assim, foram controladas as diversas possíveis leituras que os verbos poderiam ter. Abaixo, seguem-se exemplos das sentenças do *corpus* com a classificação dos verbos.⁵

Intransitivo

(14) a Maria Clara saiu hoje.

Transitivo com leitura intransitiva

Situação: a Maria Clara sofre de anorexia. A mãe, preocupada, liga para a empregada e pergunta:

Q: quando foi a última vez que a Maria Clara comeu?

(15) a Maria Clara comeu hoje, não se preocupe.

Transitivo com objeto nulo

Q: quando a Maria Clara comeu aquela torta suíça que estava na geladeira?

(16) a Maria Clara comeu hoje.

Transitivo com objeto realizado

Q: quando a Maria Clara comeu aquela torta suíça que estava na geladeira?

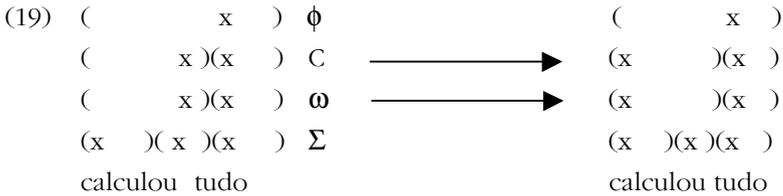
(17) a torta suíça de chocolate, a Maria Clara comeu hoje.

Predicado complexo

Q: a Maria Clara comeu feijão?

(18) a Maria Clara comeu hoje

Alguns cuidados metodológicos foram tomados para a construção das sentenças-teste. O primeiro deles diz respeito à quantidade de sílabas do verbo. De acordo com as propostas sobre regra rítmica, no caso de retração acentual, o primeiro acento do encontro move-se para a esquerda, para a primeira sílaba portadora de acento no nível inferior. Se a palavra tiver mais do que duas sílabas pré-tônicas, essa retração fará com que o acento se mova para a sílaba portadora do acento secundário, como é possível observar em (19):



Quando a regra de resolução acentual ocorre com verbos dissílabos, a retração do primeiro acento faz com que este recaia sobre a primeira sílaba da palavra, que é mais fraca, como pode ser observado em (20):⁶



Um segundo cuidado com as sentenças testadas foi que os fs formados por reestruturação tivessem o mesmo tamanho. Sandalo & Truckenbrodt (2002) argumentam que a retração acentual é bloqueada quando as frases fonológicas são de tamanho diferente, como é possível observar abaixo:⁷

(21) (CAfé QUENte) (queima a boca).

(22) # (CAfé QUENte) (queima).

5. ANÁLISE DOS DADOS

As sentenças foram ouvidas por quatro informantes. Os julgamentos sobre a possibilidade de aplicação da regra de retração acentual estão organizados na tabela em (23) de acordo com o número da sentença-teste:^{8,9}

(23)

Verbos com leitura intransitiva					Verbos transitivos com objeto realizado					Verbos transitivos com objeto nulo					Predicados complexos				
3	ap	ap	ap	ap	2	ap	ap	ap	ap	1	bk	bk	bk	bk	17	ap	ap	bk	bk
6	ap	bk	ap	ap	5	ap	ap	ap	ap	4	bk	bk	bk	bk	18	bk	ap	ap	ap
9	ap	ap	ap	ap	8	ap	ap	ap	ap	7	bk	bk	bk	bk	19	ap	ap	bk	ap
13	ap	ap	ap	ap	14	ap	ap	ap	ap	10	bk	bk	bk	bk	23	??	ap	ap	bk
20	ap	ap	ap	ap	16	bk	ap	ap	ap	11	bk	bk	bk	bk	26	ap	ap	ap	ap
21	??	ap	??	ap	24	ap	ap	ap	ap	12	bk	bk	bk	bk	27	ap	ap	ap	ap
22	bk	ap	ap	ap	25	ap	ap	ap	ap	15	bk	bk	bk	bk	28	??	ap	ap	ap

Como apresentado na seção 2, segundo Santos (2002), a retração acentual é opcional e sensível a categorias sintáticas foneticamente vazias. No caso de um encontro acentual entre verbo e advérbio, a retração do primeiro acento para a sílaba anterior é possível se a categoria vazia é um *vestígio* de movimento, mas não se é um pronome nulo (*pro*).

As sentenças aqui analisadas são sentenças simples, e a possibilidade de retração acentual em sentenças com verbos transitivos e objeto nulo indicaria que o objeto nulo se moveu para uma posição mais alta, deixando um vestígio entre o verbo e o advérbio, como em (24a). Por outro lado, a impossibilidade pode ser devida ao *pro* preenchendo a posição de objeto (24b) ou à opcionalidade da regra:

(24a) OP o José Maria contou *t* hoje.

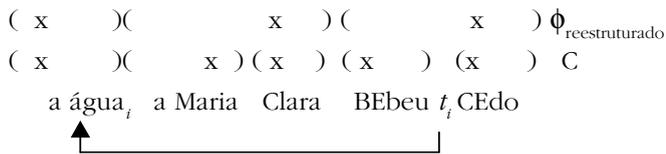
(24b) o José Maria contou *pro* hoje.

Os resultados encontrados neste trabalho apontam para o fato de que a retração é sempre bloqueada nos casos de leitura transitiva com objeto nulo em sentenças simples. Deve-se ter em mente que a regra de retração acentual é uma regra opcional, e por isso o bloqueio não é prova de que haja um *pro* entre verbo e advérbio; o fato de poder ser um vestígio e ainda assim a escolha por não se aplicar a regra deve ser considerado. Procurou-se testar as sentenças observando se fronteiras do nível entonacional ou de sentença poderiam estar influenciando a retração e, perceptualmente, tal fato não foi relevante para os informantes.¹⁰

Para todos as outras estruturas verbais a retração foi possível, mas para a leitura transitiva com objeto nulo, o bloqueio foi categórico. Se estivéssemos frente ao caso de opcionalidade da regra, este bloqueio não deveria ser categórico, deveríamos ter encontrado variação e possibilidade de aplicação da regra. Tal fato é uma evidência a favor da análise de Galves para o PB de que o objeto nulo é um *pro* não coindexado e não se move para posições mais altas da estrutura em sentenças simples.

Outro grupo de verbos testados foram os verbos com leitura transitiva e objeto realizado. Este grupo foi utilizado como controle para o grupo anterior. Os dados com verbos transitivos com objeto realizado indicam que há um movimento do objeto, deixando um vestígio. O vestígio não é enxergado pela operação de retração acentual e o encontro acentual pode ser desfeito, como é possível observar em (25):

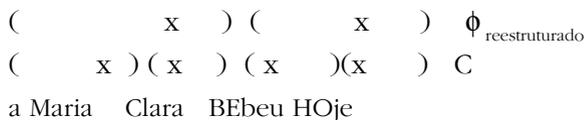
(25 (sentença-teste 2))



Uma vez que a retração acentual é opcional, é possível haver o bloqueio, como ocorre em (16). Esses resultados confirmam os encontrados por Santos (2002) para sentenças simples em que o objeto se move para a posição de tópico e deixa um vestígio na posição abaixo.

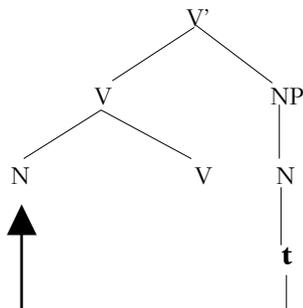
Os dados com leitura intransitiva mostram que a retração acentual é permitida, o que indica que não há nada intervindo entre o verbo e o advérbio, como em (27).

(27 (sentença-teste 3))



Nos casos de verbos intransitivos, esses dados são compatíveis com a proposta de Hale & Keyser (1993) de que os verbos intransitivos puros são formados por argumentos incorporados que deixam vestígio, como representado em (28).

(28) a Maria Clara 'deu (tomou) uma bebida' hoje. → a Maria Clara bebeu hoje.



As sentenças 17 e 18 são formadas por verbos que formam um predicado complexo, o que explica a variação nos julgamentos. “jogar bola” tem uma interpretação já lexicalizada e o nome ‘bola’ não é um complemento do verbo ‘jogar’, tanto que não é possível acrescentar um complemento ao nome ‘bola’ (ao contrário de ‘jogar a bola’, em que ‘a bola’ é complemento de ‘jogar’), como é possível observar no contraste entre (29a), um predicado complexo, e (29b), um verbo e complemento:

(29a) *jogar bola verde

(29b) jogar a bola verde

Nestes casos, por ser um predicado complexo, não há nem objeto nulo nem vestígio entre o verbo e o advérbio, e o encontro acentual pode ser desfeito, como em (30):

(30 (sentença-teste15))

Q: quando a Maria Clara jogou bola?

a Maria Clara JOgou TERça depois do almoço.

6. CONSIDERAÇÕES FINAIS

Os resultados encontrados neste trabalho corroboram certas propostas teóricas que podem ser considerados como evidências no caso de mais de uma proposta. Outro fator importante diz respeito à discussão sobre o acesso a outros componentes gramaticais.

Observou-se que, no caso de sentenças simples com objetos nulos em português brasileiro, os dados indicam que a análise de Galves, que assume que o objeto nulo em português brasileiro é um *pro in situ*, é mais acertada do que a de Raposo, que considera o objeto nulo o resultado de movimento. No que se refere a verbos transitivos com leitura intransitiva, esta análise é compatível com Hale & Keyser para verbos intransitivos. No caso de predicados complexos, os resultados são compatíveis com a interpretação lexicalizada dos mesmos.

Os resultados encontrados para os diversos tipos de objetos apontam para a distinção entre as duas categorias vazias *pro* e vestígio. Nespor & Vogel afirmam que uma versão forte da hipótese de acesso direto da fonologia à sintaxe é que todos os elementos foneticamente vazios (vestígio de NP, PRO, vestígio-*Wh*, etc) teriam o mesmo status que itens lexicais com realização fonética e bloqueariam adjacência. Neste caso, vestígio e *pro* deveriam provocar os mesmos resultados. No entanto, se assumimos a proposta de Chomsky (1995) de que vestígios são cópias apagadas, a distinção entre *pro* e vestígio torna-se clara. Vestígio passa a ser uma notação em lugar de onde foi copiada uma categoria, mas não há nada naquele lugar da estrutura. Assim, em uma sentença em que o objeto foi movido, não há nada entre o verbo e o advérbio, o que cria a adjacência acentual a ser resolvida. Essa invisibilidade de vestígios descrita acima nos leva a concluir que o apagamento de cópias se dá antes que o acento seja computado. Essa conclusão é compatível com a proposta de Nunes (1995, 1999), de acordo com a qual cópias são apagadas no componente fonológico para permitir que a estrutura sintática seja linearizada de acordo com o Axioma de Correspondência Linear de Kayne (1994).

NOTAS

¹ Apresento apenas os domínios de ϕ e C, relevantes para a discussão sobre a retração acentual.

² Para indicar a impossibilidade de retração acentual, uso #. As maiúsculas são utilizadas para marcar a sílaba acentuada.

³ Para diferentes análises da ausência de efeitos de ilha em construções com objeto nulo em PB, ver, entre outros, Kato (1993), Cyrino (1994) e Ferreira (2000).

⁴ Para a continuação deste trabalho, pretende-se fazer uma análise acústica das sentenças analisadas. A discussão sobre a realização da retração acentual realmente ocorrer ou ser uma impressão não é nova. Barbosa e Madureira, em suas apresentações neste Seminário, questionam a realização fonética da retração acentual. É claro, no entanto, para os informantes, que há um movimento acentual da primeira sílaba envolvida no encontro de acentos.

⁵ As sentenças utilizadas no teste estão no Apêndice, ao final do texto.

⁶ A literatura não é consensual sobre a formação de pés nestes casos. Propostas como a de Nespor & Vogel (1986) e Selkirk (1984) consideram que a primeira sílaba forma um pé degenerado. Sandalo (comunicação pessoal) aponta que também é possível a criação de um iambo.

⁷ Os exemplos são de Sandalo (2002). Para uma discussão pormenorizada de sua proposta, veja o texto neste mesmo volume, e Sandalo & Truckenbrodt (2002).

⁸ *ap* para possibilidade, *bk* para bloqueio, ?? para dúvidas.

⁹ As sentenças 3, 6, 9, 13 são formadas por verbos transitivos com leitura intransitiva. As sentenças 20, 21, 22 são formadas por verbos intransitivos puros.

¹⁰ Como dito em nota anterior, um próximo passo deste trabalho será fazer uma análise acústica das sentenças.

REFERÊNCIAS BIBLIOGRÁFICAS

ABOUSALAH, E. S. F. *Resolução de choques de acento no português brasileiro: elementos para uma reflexão sobre a interface fonologia-sintaxe*. 1997. Dissertação (Mestrado) – Universidade Estadual de Campinas.

CHOMSKY, N. *The minimalist program*. Cambridge, Mass: The MIT Press, 1995

CHOMSKY, N.; HALLE, M. *The sound pattern of English*. New York: Harper and Row, 1968

- CYRINO, S. *O objeto nulo no português do Brasil – um estudo sintático-diacrônico*. 1994. Tese (Doutorado) – Universidade Estadual de Campinas.
- FERREIRA, M. B. *Argumentos nulos em português brasileiro*. 2000. Dissertação (Mestrado) – Universidade Estadual de Campinas.
- FROTA, S. Clashes and prosodic domains in European Portuguese. *IFA Proceedings*, n. 19, p. 93-107, 1995.
- GALVES, C. O objeto nulo no português brasileiro: percurso de uma pesquisa. *Cadernos de Estudos Lingüísticos*, n. 17, p. 65-90, 1989.
- HALE, K.; KEYSER, S. J. On Argument structure and the lexical expression of syntactic relations. In: HALE, K.; KEYSER, S. J. (Ed.). *The view from Building 20 - essays in linguistics in honor of Sylvain Bromberger*. Cambridge, Mass: MIT Press, 1993.
- JAEGGLI, O. Remarks on *To* contraction. *Linguistic Inquiry*, n. 11, v. 1, p. 239-245, 1980.
- KAISSE, E. M. *Connected Speech – the Interaction of Syntax and Phonology*. Florida: Academic Press, Inc., 1985.
- KATO, M. A. Recontando a história das relativas em uma perspectiva paramétrica. In: ROBERTS, I.; KATO, M. A. (Org.). *Português brasileiro – uma viagem diacrônica*. Campinas: Ed. Unicamp, 1993.
- KATO, M. A.; RAPOSO, E. O objeto nulo definido no português europeu e no português brasileiro: convergências e divergências. In: CORREIA, C.; GONÇALVES, A. (Ed.). *Actas do XVI Encontro Nacional da Associação Portuguesa de Linguística APL*. Lisboa, 2001.
- KAYNE, R. *The antisymmetry of syntax*. Cambridge, Mass: The MIT Press, 1994.
- NESPOR, M.; VOGEL, I. *Prosodic phonology*. Foris: Dordrecht, 1986.
- NUNES, J. M. The copy theory of movement and linearization of chains in the Minimalist Program. 1995. Tese (Doutorado) – Universidade de Maryland, College Park.
- NUNES, J. M. Linearization of chains and phonetic realization of chain links. In: EPSTEIN, S.; HORNSTEIN, N. (Ed.). *Working minimalism*. Cambridge, Mass: The MIT Press, 1999.
- RAPOSO, E. On the null object in European Portuguese. In: JAEGGLI, O.; CORVALAN, C. S. (Ed.). *Studies in romance linguistics*. Foris: Dordrecht, 1986.

RAPOSO, E. Definite/zero alternations in Portuguese. In: SCHWEGLER, A.; TRANEL, B.; URIBE-ETXEBARRIA, M. (Ed.). *Romance Linguistics: Theoretical Perspectives*. Amsterdam, John Benjamins, 1998.

SANTOS, R. S. Categorias sintáticas vazias e retração de acento em português brasileiro. *DELTA*, n. 18, v. 1, p. 67-89, 2002.

SANDALO, F. Fonologia Prosódica e Teoria da Otimidade: reflexões sobre a interface sintaxe e fonologia na formação de sintagmas fonológicos. Comunicação apresentada no I Congresso Internacional/ VII Congresso Nacional de Fonética e Fonologia. Belo Horizonte: UFMG, 2002.

SANDALO, F.; TRUCKENBRODT, H. Some notes on phonological phrasing in Brazilian Portuguese. In: CSIRMAZ, A.; LI, Z.; NEVINS, A.; VAYSMAN, O.; WAGNER, M. (Ed.). *Phonological Answers: MIT Working Papers in Linguistics*, n. 42, p. 285-310, 2002.

SELKIRK, E. O. The role of prosodic categories in English word stress. *Linguistic Inquiry*, n. 11, p. 563-605, 1980.

SELKIRK, E. O. *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass: The MIT Press, 1984.

APÊNDICE:

Abaixo seguem-se todas as sentenças testadas com os contextos utilizados, de modo a elucidar a leitura do verbo que se pretendia testar. As sentenças estão organizadas pelos verbos utilizados, independente de sua leitura.

situação: a Maria Clara precisa beber no mínimo 2 litros de água por dia para seus rins funcionarem. A mãe chega do trabalho, pergunta e a irmã responde:

Q: Quando a MC bebeu a água dela?

- 1) a Maria Clara BEbeu CEdo, logo que você saiu
- 2) a água, a Maria Clara BEbeu CEdo, logo que você saiu.

situação: a Maria Clara tem um problema com bebidas, tem mania de se embriagar. A mãe chega de viagem, pergunta e a irmã responde:

Q: quando foi a última vez que a Maria Clara bebeu? (ser bebedora)

- 3) a Maria Clara BEbeu HOje, depois de você sair.

situação: a Maria Clara precisa tomar um remédio na hora certa, pra não ter problemas de saúde. A mãe chega do trabalho, pergunta e a irmã responde:

Q: Quando a MC bebeu o remédio que estava em cima da mesa?

- 4) a Maria Clara BEbeu HOje, depois que eu cheguei.
- 5) o remédio, a Maria Clara BEbeu HOje, quando eu saí.

situação: a Maria Clara tem câncer de pulmão. A mãe chega de viagem e pergunta, a irmã responde:

Q: quando foi a última vez que a Maria Clara fumou? (ser fumante)

- 6) a Maria Clara fumou hoje, logo antes de você chegar.

situação: a Maria Clara adora fumar cachimbo e sua mãe detesta esse hábito, pois acha que cachimbo é coisa de homem. A mãe chega de viagem, pergunta e a irmã responde:

Q: Quando foi a última vez que a MC fumou o cachimbo do pai?

- 7) a Maria Clara FUmou HOje, logo antes de você chegar.
- 8) o cachimbo do papai, a Maria Clara FUmou HOje, escondido de mim.

situação: a Maria Clara foi ao shopping para fazer várias compras. O marido foi levá-la e ficou esperando no estacionamento. Ele vira para o amigo e diz:

9) a maria clara COMprou RÁpido hoje. (fez as compras rápido)

situação: a Maria Clara compra bilhetes de loteria todos os dias, com esperança de ganhar na sena. O marido foi levá-la e ficou esperando no estacionamento. Ele vira para o amigo e diz:

- 10) a Maria Clara COMprou RÁpido hoje.

situação: um homem está conversando com seu amigo sobre as compras de natal. O amigo diz que a esposa irá fazer compras no sábado e ele responde:

11) a Maria Clara COMprou HOje mesmo, pra evitar tumulto no shopping.

situação: um homem está conversando com seu amigo sobre o carro novo que a Maria Clara quer comprar. O amigo diz que a esposa irá comprar um no sábado e ele responde:

12) a Maria Clara COMprou HOje, porque não podia ficar sem carros.

situação: um homem está conversando com seu amigo sobre as compras de natal. O amigo diz que a esposa irá fazer compras no sábado e ele responde:

13) a Maria Clara COMprou CEdo este ano, antes de todo mundo.

situação: um homem está conversando com seu amigo sobre um boné comemorativo da seleção. O amigo diz que não conseguiu comprar, pois estavam todos vendidos. O marido diz:

14) a Maria Clara COMprou CEdo o boné

situação: a Maria Clara precisa comprar um par de sapatos novos para ir a uma festa. Mas ela é conhecida por ser muito indecisa e demorar uma eternidade para escolher o mais bonito. A mãe pergunta e a irmã responde:

Q: Como a Maria Clara comprou os sapatos desta vez?

15) a Maria Clara COMprou RÁpido.

16) os sapatos, até que a Maria Clara COMprou RÁpido dessa vez.

situação: a Maria Clara adora esportes e pratica vários. A mãe pergunta e a irmã responde:

Q: Quando a Maria Clara jogou bola?

17) a Maria Clara JOgou TERça depois do almoço.

18) bola, a Maria Clara JOgou TERça depois do almoço.

situação: jogar o lixo é uma tarefa das crianças, que se revezam. A mãe pergunta e o pai responde:

Q: hoje é dia de quem colocar o lixo na rua?

19) a Maria Clara jogou ontem, hoje é a vez do João.

Situação: a Maria Clara e a Joana estão de castigo, proibidas de sair de casa. A mãe vai a um congresso e quando volta, pergunta e a empregada responde:

Q: alguém me desobedeceu?

20) a Maria Clara saiu ontem.

Situação: a mãe viaja para um congresso e deixa a filha pequena com a sogra. Quando volta pergunta e a sogra responde:

Q: a Maria Clara deu trabalho?

21) a Maria Clara chorou ontem quando você saiu, mas depois se acostumou.

Situação: há dois congressistas hospedados num hotel. O coordenador do evento, preocupado, liga para um amigo e pergunta se eles já foram, e o amigo responde:

Q: alguém já foi embora?

22) a Maria Clara partiu ontem, mas o João Pedro só vai amanhã.

Situação: a Maria Clara e a Joana não gostam de comer feijão. A mãe chega do trabalho, pergunta e a empregada responde:

Q: alguma delas comeu feijão?

23) a Maria Clara comeu hoje, mas a Joana não quis nem saber.

Situação: a Maria Clara e a Joana têm que comer feijão e fígado todos os dias por causa de uma anemia. A mãe esteve viajando. Quando volta, pergunta e a empregada responde:

Q: elas comeram direitinho?

24) o fígado, a Maria Clara comeu hoje no almoço, mas feijão não teve jeito.

Situação: a Maria Clara e seus amigos estão acampando. Eles lhe arranjam um violão e pedem para ela cantar; o namorado responde.

Q: canta uma música pra gente, Maria Clara.

25) Andança, a Maria Clara cantou ontem. Não aguento mais.

26) a Maria Clara cantou ontem à noite, hoje é a vez do João.

Q: toca violão pra gente, Maria Clara.

27) a Maria Clara tocou ontem, hoje ela vai descansar.

Situação: seu Geraldo adora jogar bingo com a família mas não pôde ir à última rodada. Ele pergunta e a esposa responde.

Q: algum de nós ganhou alguma coisa?

28) a Maria Clara jogou ontem, mas não ganhou nada.

Da frase fonológica no dialeto goiano*

Irene Zasimowicz Pinto Calaça**
Universidade Federal de Minas Gerais

Abstract

We present a description of the phonological phrase in a variety of Brazilian Portuguese spoken in the State of Goiás, according mainly to Nespór and Volgel's theory of constituents (1994). Other sources are Mira Mateus (1990) and Tcheremíssina (1969).

We intend to demonstrate that it is important to consider some prosodic elements, such as pauses, rhythm, stress, time and sandhi in order to identify phonological phrase in a more adequate way.

Neste trabalho apresentamos uma breve descrição da frase fonológica sob o ponto de vista prosódico, tendo por base uma das variedades goianas. Assim, numa pequena introdução, apresentamos ao leitor os constituintes prosódicos de Nespor e Vogel (1994), bem como a definição teórica da frase fonológica, para então discorrermos sobre os elementos prosódicos que se destacaram ora na reunião, ora na segmentação das FFs (pausas, acento e ritmo, tempo, sândi e juntura). Todos os dados foram retirados de nossa pesquisa de mestrado (CALAÇA, 1999)

1. A TEORIA DOS CONSTITUINTES PROSÓDICOS

A teoria dos constituintes prosódicos de Nespor e Vogel (1994) distingue sete níveis ou hierarquias prosódicas, as quais, num crescendo, seriam: *sílaba* (δ), *pé* (Σ), *palavra fonológica* (ω), *grupo clítico* (C), *frase fonológica* (Φ), *frase entonacional* (I) e *enunciado* (U). Esses constituintes são unidades lingüísticas complexas, compostas através da dominância de um termo, o núcleo, considerado forte, sobre outro(s) mais fraco(s). As unidades inferiores encontram-se totalmente inseridas nas de hierarquia superior.

De acordo com as autoras, (1994: 13-18), os constituintes prosódicos situam-se no nível em que processamos a percepção da fala e estão interrelacionados a diversos outros níveis gramaticais (como morfológico, sintático, semântico...), sem, no entanto, se confundirem com nenhum deles nem tampouco se inserirem completamente nos mesmos.

Este sistema – total ou parcialmente – é encontrado em diversas línguas, e a sua existência é confirmada por regras fonológicas e de

projeção.¹ Tais regras atuam exclusivamente no nível de cada um de seus integrantes e preenchem “vácuos” que não podem ser explicados por constituintes morfológicos e sintáticos.

2. A FRASE FONOLÓGICA: TEORIA

A frase fonológica (FF) ou sintagma fonológico é um constituinte de nível médio que, de acordo com Bisol (1996: 254), seria “o constituinte que congrega um ou mais grupos clíticos, ou seja, o grupo clítico propriamente dito e a palavra morfológica, ambos **C** neste nível.” O grupo clítico (**C**) seria uma forma semelhante a uma palavra que, destituída de acento, não mantém sua independência total em um enunciado, ligando-se, então, a uma palavra fonológica, por exemplo, as expressões: “de vidro”, “de relance”. Assim, de acordo com a definição supra, a frase fonológica seria constituída por um ou mais grupos clíticos.

Seguindo fielmente a linha de Nespor e Vogel (1994), a lingüista brasileira explica-nos que a frase fonológica pode ser definida a partir dos três princípios a seguir:

- (I) o domínio da FF abrangeria o **C** que contém o núcleo lexical (X) e todos os demais **Cs** de seu lado recursivo, até o próximo núcleo lexical (X₂), fora da projeção máxima de (X);
- (II) construir-se-ia a FF agregando-se, de forma n-ária, todos os **Cs** inclusos na cadeia supra mencionada;
- (III) a proeminência relativa dentro da cadeia n-ária seria determinada pela ramificação sintática da língua em questão. Assim, em línguas ramificadas à direita, como o português, o nó mais forte (“s”) – também denominado por Bisol (1996: 255) como “cabeça de constituinte prosódico” – seria o mais à direita; em línguas cuja ramificação faz-se à esquerda, o nó localizado mais à esquerda seria rotulado de “s”. Os demais nós, irmãos de “s”, são rotulados de fracos (“w”).

Nenhuma das autoras mencionadas especifica com maiores detalhes, o que faz com que um elemento que se sobressai, seja

proeminente em relação aos demais. A esse respeito optam Nespor e Vogel (1994: 19) por “*uma convenção mais simples*”, *ipsis litteris*, a saber: “[...] a relação de proeminência relativa entre os nós irmãos se caracteriza pela atribuição do valor forte (s) a um dos nós e o valor (w) ao restante.”

Contudo, é possível inferir que o termo “proeminência relativa” também estaria ligado a questões de tonicidade, já que, páginas adiante, mencionam (Idem, *ibidem*: 23):

No que se refere a detalhes dos esquemas de proeminência relativa, [...] sem dúvidas, será necessário um sistema mais complexo, como, por exemplo, aquele que combine as propriedades da grade [métrica] com os âmbitos da hierarquia prosódica... [nossa sublinha]

Conclui-se, assim, que na frase fonológica teria lugar um núcleo lexical (denominado (X)), que seria determinado tanto pela posição sintática “*mais à direita*” como pela tonicidade (vide nossa sublinha na citação supra).

Segundo Nespor e Vogel (1994: 195-196), o papel de núcleo lexical (X) pode ser exercido por nomes (substantivos e adjetivos) e verbos, porém a categoria do elemento não é mais relevante do que a posição que ocupa dentro de FF. Assim, todos que se encontram no lado recursivo de X, dentro de sua projeção máxima – ainda que seja outra categoria principal – se inscrevem na mesma FF. Categorias sintáticas principais são núcleos prosódicos só quando estão em posição não-marcada. Por exemplo, quanto à colocação de adjetivos e substantivos, as línguas românicas consideram como *posição não-marcada* o adjetivo que segue o substantivo, e como *posição marcada* o adjetivo antecedendo o substantivo. No segundo caso (posição marcada – adjetivo antecedendo substantivo) ambos os componentes seriam agrupados em uma única FF, o que não ocorreria na outra posição (não-marcada). Visualizemos o afirmado, que será comentado adiante.

(1) posição não-marcada: | menino |_{FF} | bom |_{FF}

(2) posição marcada: | bom menino |_{FF}

3. A FRASE FONOLÓGICA: PRÁTICA

No presente trabalho analisamos dados obtidos a partir de gravações de conversas espontâneas e de leituras de textos² por informantes experientes goianos. Ao aplicarmos a teoria acima exposta, encontramos certa dificuldade, pois ela não permitiu uma delimitação natural da FF. Embora tratem de um constituinte *prosódico*, Nespor e Vogel (1994) não desvinculam seu trabalho da base sintática, em torno da qual versou a maior parte das explicações acerca da FF. É certo que mencionam superficialmente a questão da tonicidade, porém esta não foi trabalhada mais detalhadamente – e o resultado dessa omissão pode ser visualizada no seguinte trecho retirado de Frota (1994: 82), que aplica a teoria *ipsis litteris* ao português europeu:

(3)“...[As alunas] ϕ [todas] ϕ [ofereceram-lhe] ϕ [rosas] ϕ ”

As divisões efetuadas em (3) não lembram frases fonológicas, mas simples clíticos isolados dispostos lado a lado. Onde estaria a dominância do núcleo sobre os demais? Como determinar que nome (substantivo ou adjetivo) ou verbo seria o núcleo de FF em sintagmas expandidos, sem segmentá-lo em “frações clíticas”?

Ao estudarmos enunciados, verificamos que, dentre os elementos prosódicos, mostraram-se mais relevantes na delimitação das FF as pausas, ritmo, acento, tempo, sândi e juntura, os quais examinaremos brevemente a seguir:

3.1. Pausas

As pausas – pequenas interrupções na seqüência sonora do enunciado que limitam os constituintes – têm as mais diversas funções: atuar aerodinamicamente, permitindo que o falante respire durante a fala; servir a este como instrumento de segmentação sintática ou como indicativo de alterações bruscas no enunciado, entre outras. Conforme Mira Mateus (1990: 203), “As pausas coincidem, de um modo geral, com fronteiras sintáticas, marcando freqüentemente [sic] as fronteiras de frase e, no interior das frases, as fronteiras de constituinte.”

A autora também reconhece a existência de “pausas potenciais”, ou seja, das variações melódicas que podem ser atribuídas a um mesmo enunciado, sem que os limites entre os constituintes sejam delimitados por pausa real. Estas pausas “potenciais” podem vir a ser localizadas por meio de instrumentos, através do registro da frequência fundamental (F_0). Mira Mateus afirma (1990: 203):

Freqüentemente, a frequência [sic] fundamental desce sobre a última sílaba acentuada de uma unidade entonacional e volta a subir até a primeira sílaba acentuada da unidade seguinte, sem que tal movimento seja acompanhado de pausa.

Tais variações melódicas (tonais) do enunciado encontram-se em posições, nas quais virtualmente poderia instalar-se uma pausa.

Khavrônina e Krylova (1989: 16-17) parecem ter o mesmo posicionamento de Mira Mateus, ao afirmarem que, entre tópico e comentário, pode não vir a ocorrer pausa, mas uma “sensação” de pausa, originada pela alteração de tons. Acreditamos que tal “sensação de pausa” corresponda à “pausa virtual”.

A questão da pausa virtual surgiu em nossa pesquisa, pelo fato de termos selecionado alguns textos sem vírgula, nos quais os nomes foram expandidos com vários determinantes, o que originou sintagmas bastante longos. Os informantes, contudo, segmentaram sem dificuldade essas passagens, utilizando-se ora de pausas reais, ora de pausas virtuais – essas últimas menores e ligadas a fatores supra-segmentais. Cagliari (1992: 143) está entre os autores que discorda da noção de pausa virtual.

Em nosso trabalho constatamos a grande importância de pausas – reais e virtuais – na delimitação da FF. As pausas reais indicam, quase sempre, o término de uma frase entonacional ou de um enunciado e o início do próximo. Como os constituintes inserem-se totalmente uns nos outros, logicamente, a fronteira de uma pausa real apontaria o término de alguma FF, inserida nos outros constituintes de hierarquia superior, como num encaixe sistemático.

A realização de uma pausa e sua duração dependem, por um lado, das necessidades respiratórias do falante e, por outro, das

funções sintáticas que exercem. Em pausas reais, essas funções (segmentação sintática, alterações de enunciado, entre outras) ligam-se a domínios superiores à FF e são de vital importância, pois delas dependem a continuidade melódica do enunciado e sua coesão. Diante de uma pausa real manifestam-se certas características de cunho segmental, como a não-assimilação de contóides surdos diante de vogais, como em:

(4) “em que o rema coincide com certos papéis semânticos||] em presença de certos outros.”

“Aí convidou ele[s||] aí eles sentaram.”

A conservação da qualidade desses contóides não costuma ocorrer no interior da FF, no qual o sândi externo dissolve antigas fronteiras entre palavras e origina a assimilação por sonoridade de contóides surdos:

(5) “tod[uzuze]nunciados...”

“Do jui[z'ε:]

“Tod[uzuze]lefantes...”

A assimilação por sonoridade seria, então, uma regra de interior de domínio da frase fonológica.

Pausas virtuais são obtidas através de valores relativos, porém podem ser registradas fisicamente, ao compararmos altura tonal ou mesmo alternâncias de velocidade de um determinado segmento com tom e velocidade de outro. Se após um *continuum* alto, ascendente, ou mesmo descendente, temos imediatamente um tom médio, cria-se uma situação de pausa virtual.

A pausa virtual é realçada ainda mais pelas alterações de velocidade do informante:

(6) “É por causa do cursinho, Irene. Cê vê, o cursinho é assim...”

[,ε'p̃,ka'wzε'dukur'si:ŋu | ɪ,rē'ŋɪ || ,sε've: | ukur'si:ŋu || ,ε'e'si:||]

No exemplo acima, observa-se o destaque semântico dado a certos vocábulos em função da simples alteração de velocidade, o qual auxilia, inclusive, na manutenção da coesão do enunciado.

Ao nosso ver, pausas virtuais manifestam plenamente os contornos de FF, podendo ser visualizadas, em primeiro lugar, pela variação do *tom*, não obstante variações tonais também ocorrerem em instâncias maiores, como **I**, sob a forma de tessituras – isto é, palavras e frases inseridas no discurso principal. As diferenças entre um e outro empregos seriam funcionais e físicas: em **I**, a elevação ou abaixamento tonal (juntamente com a entonação) poderiam estender-se por um âmbito maior do que o da FF. Acreditamos que (7) ilustre o afirmado:

(7) “A mesma tradição de análise a que se deve o conceito de articulação marcada...”

[e'me:zmø⁺ | t̥rɛd^j,sẽw̃d^jje'na:lɪzɪ]_I [ɛcɪsɪ,d̥ɛ'vju^kõ'se:jt̥u|
d^jɪ, ɛ'rt̥^jɪku^le'sẽ:w̃| mɛx'ka:d̥ø⁺||]_I



(8) de articulação | marcada

Em (7), a tessitura pertencente a segunda I inicia-se em “...a que se deve...” e encerra-se em “...marcada”. Dentro desse período levado em tom baixo, encontramos três frases fonológicas com seus respectivos núcleos acentuais e contornos. Por outro lado, em FF, a alteração tonal de que nos valemos restringe-se ao limite entre palavras – a micromelodia daquela mais à direita, seguindo seu curso normal dentro do constituinte, como em (8).

Numa segunda instância, quando não é possível delimitar a FF unicamente pelo tom, devemos observar a *velocidade do enunciado*, que se torna mais *lenta* durante a pronúncia do núcleo de FF. Caso ambos os fatores anteriores estejam neutralizados, vemos a força acentual influir na cisão de FF; afinal, em cada frase fonológica temos um único núcleo, que se sobressai perante os demais componentes.

Segundo Mira Mateus (1990: 203), normalmente não ocorrem pausas entre especificadores e o núcleo sintático. Somente em casos de ênfase, quando então “...qualquer elemento pode constituir, por si só, uma unidade entonacional, incluindo um clítico.”

3.2. Acento e Ritmo

Verificamos que os falantes manobram a tonicidade criando uma hierarquia de acentos, com o intuito de manter a unidade e a coerência semântica de FF. Como isso se explicaria? Isoladamente, a palavra é constituída de sílabas átonas agrupadas próximas à sílaba tônica, portadora de acento – este último, um elemento estruturante. Na cadeia da fala, porém, a palavra passa a integrar novos e complexos níveis, perdendo parcialmente sua autonomia. Os acentos tônicos individuais de cada palavra passam a ser secundários em relação ao acento principal da frase fonológica. Portanto, em cada FF encontramos um único acento principal, um único núcleo tônico, que parcial ou totalmente é pronunciado com maior intensidade e lentidão do que os demais integrantes do sintagma. Acreditamos ser este o principal ponto que diferencia a FF dos outros constituintes prosódicos e que nos distancia um pouco da teoria exposta por Nespor e Vogel (1994) e Bisol (1996). Eis alguns exemplos retirados de trabalho anterior (Calaça, 1999: 56-59), aos quais aplicamos a grade métrica, que corroboram o afirmado. Comparem-nos com (1) e (2):

(9) posição não-marcada:

```

                *
                *
            *    *
|os cozinheiros gordos|

```

```

                *
                *
            *    *
|a lua pálida|

```

```

                *
                *
            *    *    *
|é um edifício belo|

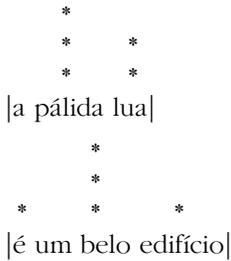
```

(10) posição marcada:

```

                *
            *    *
            *    *
|os gordos cozinheiros|

```



Eis nossa interpretao para os exemplos supra: o portugus  uma lngua com rvore sinttica normalmente ramificada  direita. Dessa forma, em enunciados no-marcados, a ltima palavra  direita seria o no mais forte (vide (9)). Contudo, enunciados estilisticamente marcados (conforme (10)) so encontrados em grande nmero, como quando adjetivos antecedem substantivos. A esse respeito, afirma Lapa (1991: 106):

...o adjetivo anteposto ao substantivo forma com ele uma espcie de grupo fraseolgico, em que ambos elementos perdem um pouco do seu valor, em proveito do conjunto.

Aparentemente, esse “grupo fraseolgico” ocorre pelo fato de intuirmos a posio pospositiva do adjetivo. Encontrando-se o mesmo deslocado, buscaremos salient-lo entonacionalmente. Porm, como a ltima palavra  direita – no caso, o substantivo – atrai nossa ateno mecanicamente, sentir-nos-emos divididos, o que pode gerar flutuaes: hierarquias acentuais dentro de FF ou mesmo sua ciso em duas FFs distintas, conforme os exemplos que seguem:

- ```

 *
 * *
 * *
(11) os gordos cozinheiros
 * *
 * *
 * *
(12) bastante veiculad | conceito
 *
 *
 * *
(13) veiculad | conceito

```

A presença de intensificadores, quantificadores e de negação também ocasiona alterações dentro de FF. A força semântica e estilística de alguns modificadores como “certo(s)”, “todo(s)”, “bastante”, “não”... pode gerar flutuações na posição do núcleo tônico:

- \*  
\*   \*  
\*   \*
- (14) os elefantes todos  
\*  
\*   \*
- (15) os elefantes todos  
\*  
\*   \*
- (16) não gosta  
\*  
\*   \*
- (17) ela não quer seguir (é nada)  
\*   \*   \*   \*

Porém, na maioria das vezes, os modificadores – pronomes demonstrativos, indefinidos e de negação como “próprio” “tudo”, “certo”, “nada”, “ninguém”, “não”... tomam para si o acento nuclear de FF.

Um estudo preliminar de Bessa Neto (1996: 162) sobre os pronomes “próprio”, “certo” e “qualquer”, por exemplo, revelou que:

...esses itens possuem boa capacidade de se mover dentro do SN [Sintagma Nominal] e que, nas diferentes posições em que podem figurar, revelam traços semânticos diferentes.

Daí resulta que as flutuações acentuais encontradas em (11)-(17) poderiam ter sido geradas por idiossincrasias de natureza semântica, em função da posição em que se encontram e da interpretação específica que o leitor/falante lhe atribuiu no momento.

Por outro lado, se formos examinar melhor os exemplos (11), (12), (14) e (16) veremos que estes não se inserem na teoria de Nespor e Vogel (1994), de acordo com a qual o nó da direita é sempre o núcleo acentual da FF (conforme exemplos ilustrativos (1)

e (2)). Aqui consideramos não apenas a base sintática, que nos pareceu muito restrita, mas também – e principalmente – a hierarquia tônica, que nos mostrou ser mais condizente com a realidade. Da mesma forma em que num vocábulo encontramos sílabas pretônica e postônicas, postulamos que, na frase fonológica, é possível encontrar clíticos átonos antes ou após o núcleo tônico, neste caso, encerrando-a e dando-lhe fluidez. O procedimento pelo qual optamos parece se adequar melhor aos dados, permitindo uma descrição mais simples e natural dos mesmos.

### 3.3. Tempo

Tempo seria a duração do enunciado, também relacionada à velocidade da fala<sup>3</sup>. É analisável a partir das variações de velocidade do “débito” – velocidade peculiar de cada falante na corrente da fala. Ou seja, não obstante os indivíduos terem débitos diferentes (cada um ter sua velocidade própria ao falar), é possível registrar seus tempos de maneira uniforme. Mira Mateus (1990: 209) completa:

É também fundamentalmente o tempo que permite distinguir, de um ponto de vista prosódico, uma frase imperativa de uma interrogativa QU-, ou mesmo de uma declarativa...

De acordo com Tcheremíssina (1982: 114-116), a organização temporal da frase fonológica depende do papel semântico das palavras que a compõem. Seu centro é sempre pronunciado mais lentamente, podendo o ouvinte discernir inclusive as mais minuciosas variações temporais do enunciado. Nos textos que examinamos, os informantes armaram uma espécie de jogo, conciliando os tempos “rápido”, “normal” e “lento” (a ênfase é marcada com sublinhado duplo) para transmitir a hierarquia semântica dentro do enunciado, como em:

(18) O irmão da menina... mas Osnilda, uma educação... você precisa de vê.

[wɪx'mã:vw̃demĩ,ŋĩ'nə' || ma'jz'hi:wðə' | ,ũ:mə'eduka'sê:w̃ | ,se'prɪ'sizə' | d'i'vê: ||]

Os trechos pronunciados mais rapidamente demonstraram que o pensamento ali expresso era secundário, formando apenas um “pano de fundo” para a linha primária, que seguiria um pouco adiante. Este corte na linha do pensamento foi salientado, às vezes, pelo próprio tom – mais baixo que o do restante da frase.

Da mesma forma, palavras e sílabas tônicas – principalmente as ligadas ao núcleo tônico da FF – eram quase sempre pronunciadas mais lentamente, propiciando um realce semântico para o trecho em questão e prendendo a atenção do ouvinte.

Constatou-se, também, que a velocidade de enunciação das frases fonológicas interfere na sua dimensão, pois as mesmas, ao serem pronunciadas mais rapidamente, ficavam, por vezes, maiores.

(19) “consiste num sistema...” Informante “c” [ | kō'si:stʰi |, ŋũ'sis'te:məʔ | ]

Informante “d” [ | kō,si:stʰi,ŋũ'sis'te:məʔ | ]

Embora maior, a frase fonológica pronunciada pelo informante “d” continua tendo apenas um núcleo acentual.

### 3.4. Sândi ou juntura

Neste trabalho, consideraremos apenas o sândi ou juntura externa, ou seja, as alterações fonético-fonológicas que ocorrem em fronteiras vocabulares, quando da justaposição de “formas livres mínimas” na corrente da fala.<sup>4</sup>

Câmara Jr. (1977: 47) explica que encontramos na palavra “... uma gradação de intensidades menores, ou atonicidades variáveis, que são foneticamente bem depreensíveis...” e que se constituem fatores determinantes na instauração dos limites vocabulares. A começar pelo acento tônico, que, num grupo de força, diminui de intensidade, subordinando-se ao último acento tônico do grupo. A subtonicidade é uma característica também encontrada em vocábulos compostos, segundo Câmara Jr., e, ao nosso ver, em frases fonológicas.

Em relação às demais sílabas, o autor registrou três graus de atonicidade nos paroxítonos (Idem ibidem, 48): a máxima “em sílabas

átonas finais”; a média “nas pretônicas não-iniciais ou iniciais começadas por vogal” e a “...atonicidade mínima, nas pretônicas iniciais começadas por consoante.”. Em vocábulos proparoxítonos as sílabas postônicas seriam ambas débeis, possuindo a atonicidade máxima.

Embora seja mecanicamente condicionado por sua posição em relação à sílaba tônica, o grau de atonicidade distingue os limites entre vocábulos dos de outras combinações semelhantes, pertencentes ao domínio interno da palavra. A atonicidade máxima seria, portanto, um indício de juntura vocabular.

Arrolaremos, a seguir, considerações sobre junções de palavras peculiares à FF e ao dialeto goiano em questão, considerando sempre que variações de registro influem nos resultados.

### 3.4.1. Junções de palavras em fronteiras da FF

- Quando dois vocóides iguais ou semelhantes encontravam-se em zonas limítrofes da FF, surgia um hiato, ou seja, manteve-se a seqüência de dois núcleos silábicos, dois tempos articulatórios. Exemplos:

(20) [u | u]

era gostos [u | u] salgadinho

pena que tava fri [u | u] salgadinho

resum[u | u]s exercício(s)

(21) [u | ũ]

eles (a)rrumar[u | ũ]ma...

(22) [i | i]

quem que é pa..crent[i | i]quem que é católico ali?

O mesmo tipo de hiato ocorreu com o contóide [s] em fronteira de FF:

(23) [s | s]

certo[s | s]lingulares...

papéi[s | s]lemânticos

- Como alguns vocóides postônicos finais estavam bastante debilitados, a sua articulação deu-se de forma incompleta, ensurdecida:

(24) [ɹ] –tradição de análise[ɹ | ]

(25) [ʊ] –(rema) referid [ʊ | ]

- fonema /s/ diante de pausas manteve suas qualidades originais, manifestando-se como o fone [s]

(26) era nove hora[ s||]

### 3.4.2. Junções de palavras no interior da FF

- O fonema /s/ entre duas palavras pertencentes a uma mesma frase fonológica realiza-se como [s] diante de fones surdos e como [z] diante de vocóides ou fones sonoros.

(27) “nós chegamos” [ |,nɔːjs[e'gõ:mu|]

(28) “e os meninos” [ |ju :zmĩñi:ɲu|]

- É comum tanto vocóides quanto contóides elidirem-se nos sândis dentro da FF:

(29) a par [tʰ] do juiz

(30) É [pʰ] causa ...

- No que tange à fala espontânea, verificamos interessante processo morfossintático em andamento, envolvendo o número plural. Nomes semanticamente plurais acompanhados de pré-determinantes plurais, não recebem a desinência que lhes compete. É assim que encontramos com grande freqüência:

(31) as minhas prova; os exercício

(32) os menino

Em leituras orais – cuidadosas e tensas – este processo é mais raro, porém conseguimos registrá-lo:

(33) tod[u] enunciad[u s]

Aqui, um pequeno comentário. Aparentemente, ao resguardar o ritmo acentual que lhe é tão característico, o falante goiano busca evitar “tautologias” numéricas, ou seja, indica número plural nos determinantes e não o repete nos nomes. Um detalhe: observamos, que, havendo dois pré-determinantes na variante oral culta ambos recebem o morfema plural. Somente o nome mantêm-se no singular, como vemos em (31). Isso gera frases fonológicas bem mais coesas, pois o nome passa a depender parcialmente de um elemento externo – o morfema plural de seu determinante – para completar-se semanticamente.

- Da mesma forma, ora os verbos têm suas desinências plurais elididas:

(34) “nóis chegam[u] cedo”; eles começ[u]  
 ....ora são mantidos na terceira pessoa do singular:

(35) era nove horas

(36) as minhas prova | foi na quarta-feira

- Nos verbos também ocorre a monotongação dos ditongos “-ou” e “õw”, que perdem o glide e alongam-se:

(37) ficou → [fr'ko:]

(38) estou → [t'o:]

(39) sentaram → [sẽ'ta:ru]

- Eliminação do “-r” em infinitivos verbais, com posterior alongamento da vogal que o antecedia:

(40) seguir → [sɨ'ji:]; contar → [kõ'ta:]

(41) tomar → [t'o'ma:]

(42) encontrar → [ĩkõ'tra:]

Desaparecimento de sílabas átonas iniciais (aférese):

(43) está → [t'a:]

(44) arrumaram → [χũ'ma:ru]

As subtrações de elementos de que aqui tratamos, bem como outras – mais esporádicas e menos sistemáticas – influem na dimensão da frase fonológica que, no falar goiano, pode vir a ser bem ampla. Sua coesão, todavia, é mantida pela hierarquização tônica e por outros fatores analisados anteriormente.

Em conclusão, um pequeno comentário: ainda no início deste capítulo, mencionamos a segmentação de um enunciado encontrado em Frota (1994: 82) e reproduzido sob o número (3), o qual nos causou estranheza. Em nosso material encontramos algo semelhante. Uma de nossas informantes segmentou as passagens: “todos interessantes enunciados” e “com certos singulares papéis semânticos” num ritmo mais silábico, que foi assimilado no contexto com um colorido semântico entonacional. Isso foi possível, pois a informante criou uma hierarquia de acentos dentro das referidas passagens. Essa hierarquia manteve a coesão de todo e auxiliou o ouvinte a assimilá-lo. Todavia, em Frota, a hierarquia acentual não foi registrada, dificultando a apreensão do conteúdo semântico. Cremos poder, inclusive, afirmar que a clitização de frases fonológicas: 1) é possível como meio estilístico, pois foge um pouco à naturalidade da fala coloquial; e 2) eleva FF a categorias prosódicas superiores, mais ligadas à questão tonal, a fim de ser bem assimilada semanticamente.

A realização do presente trabalho despertou-nos uma série de indagações acerca da formalização da lógica interna na estruturação da frase fonológica, algumas das quais esperamos solucionar durante o trabalho que ora realizamos no curso de doutorado.

## NOTAS

\*Artigo elaborado a partir da dissertação de Mestrado “A Frase Fonológica numa Variedade Lingüística Goiana”, orientada pela Profa. Dra. Marita Pôrto Cavalcante e defendida em 1999, na Faculdade de Letras da Universidade Federal de Goiás.

\*\* Mestre em Letras e Lingüística (UFG) e em Filologia Eslávica (Patrice Lumumba), Doutoranda da UFMG, bolsista da FAPEMIG.

<sup>1</sup> As regras de projeção viabilizam a interação da fonologia com os demais níveis gramaticais.

<sup>2</sup> Em textos escritos o sintagma pode ser expandido consideravelmente através de determinantes, o que permite um estudo bem mais dirigido.

<sup>3</sup> Mira Mateus (1990: 209) previne que não devemos confundir tempo e ritmo, “*na medida em que um mesmo ritmo se pode observar em tempos diferentes e vice-versa*”.

<sup>4</sup> O sândi, contudo, é um fenômeno de noções bem mais amplas, podendo ocorrer, inclusive, no interior de palavras

## REFERÊNCIAS BIBLIOGRÁFICAS:

BESSA NETO, R. Próprio, certo e qualquer. *Revista de Estudos da Linguagem*. Belo Horizonte, ano 5, p.161-170, jul./dez., 1996. (número especial).

BISOL, L. Sândi vocálico externo: degeminação e elisão. *Cadernos de Estudos Linguísticos*, Campinas, n. 23, p.83-101, jul./dez., 1992.

BISOL, L. Constituintes prosódicos. In: BISOL, L. (Org.). *Introdução a estudos de Fonologia do Português Brasileiro*. Porto Alegre: EDIPUCRS, 1996.

CAGLIARI, L. C. Prosódia: algumas funções dos supra-segmentos. *Cadernos de Estudos Linguísticos*. Campinas, n.23, p.137-151, jul/dez.,1992.

CALAÇA, I. Z. P. *A frase fonológica numa variedade linguística goiana*. 1999. Dissertação (Mestrado em Letras e Linguística) - UFG, Goiânia.

FROTA, S. Aspectos da prosódia e do foco no Português Europeu. *Letras de Hoje*, Porto Alegre, v. 29, n. 4, p.77-99, dez., 1994.

KHAVRÔNINA, S. A.; KRYLOVA, O. A. *Obutchenie inostrantsev poriadku slov v russkom yazyke*[O ensino da ordem das palavras a alunos estrangeiros de russo]. Moscou: “Russkii Yazik”, 1989.

LAPA, M. R. *Estilística da Língua Portuguesa*. 3. ed. São Paulo: Martins Fontes, 1991.

MIRA MATEUS, M. H. *Fonética, fonologia e Morfologia do português*. Lisboa: Universidade Aberta, p. 191-217, 1990.

NESPOR, M.; VOGEL, I. *La prosódia*. Tradução (primeira edição em inglês: 1986). Madrid: Visor Distribuciones, 1994. (em espanhol)

TCHEREMÍSSINA, N. V. Stroenie sintagmy v russkoi hudogestvenoi retch [ A estrutura do sintagma do discurso literário russo]. In: TCHEREMÍSSINA, N.V. (Coord.). *Sintaksis I intonatsia*. [Sintaxe e entonação]. [s.l.]: “UFA”, 1969. p. 3-95.

# **Diminutive reduplication in government phonology**

Renée Lambert  
Laboratoire Dynamique du Langage  
Université Lumière Lyon 2  
Institut des Sciences de l'Homme  
France

## **Resumo**

As propriedades morfológicas e fonológicas da formação diminutiva por reduplicação são estudadas em função das hipóteses de Lowenstamm (1996, 1999) sobre a organização universal dos segmentos fonológicos e mais particularmente sobre a sílaba inicial das palavras de categorias lexicais essenciais. É proposto que a posição inicial seja o lugar da reduplicação diminutiva em francês. Fatos do francês de Quebec bem como do antigo grego são apresentados seguidamente para apoiar a hipótese segundo a qual as palavras são representadas universalmente por uma seqüência estrita de consoantes e de vogais (CVCV...).

This article deals with the diminutive formation by reduplication in French. Those reduplications are typically formed by the copy of the first consonant and the following vowel at the left of the base, for example *filille* 'little girl', diminutive of *fille* 'girl', *gueguerre* 'squabble' (figurative) diminutive of *guerre* 'war'. The morphological and phonological properties of those formations will be studied following the hypotheses of Lowenstamm (1996, 1999) on the phonological constitution of words from major lexical categories (nouns, verbs). It will be shown that the properties of the diminutive formation in French are compatible with the three next statements: 1) the initial CV site proposed by Lowenstamm (1999) is the place of diminutive reduplications in French; 2) this type of reduplications is restricted to the size of the initial site, that is the first consonant and vowel only; and 3) this process of reduplication is productive. Those hypotheses make predictions, first, that nouns beginning with an initial consonant cluster won't reduplicate, and second, that nouns semantically compatible with the diminutive interpretation will be able to form diminutive by reduplication. To assert those propositions, on one hand, we will present different types of diminutives obtained by reduplication in French and Canadian French, in order to show that this process is productive and that effectively consonant clusters do not reduplicate to form diminutives (at least, not with apparent implication of the initial site), and on the other hand, data from Canadian French and Ancient Greek will be analyzed to show that the universal organization of segments must be a strict consecution of consonants and vowels.

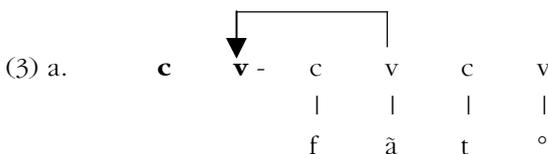
In the first section, we present the theory of government phonology and its new developments. The second section will be devoted to the morphological properties of the diminutive reduplication.

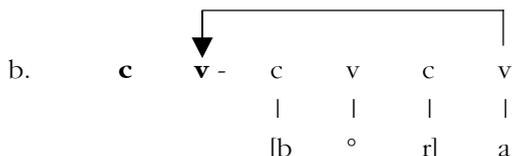


Furthermore, what has intrigued Lowenstamm in his 1999 paper are languages like French and English with no templatic phonology. Because of the proper government defined between vocalic constituents only, the prediction is that there will be no restrictions concerning consonants. Nevertheless, in French for example, we can only observe the sequences occlusive-liquid, but never the reverse one, occlusive-liquid (pl, \*lp). Moreover, no language presents the sequence of decreasing sonority without presenting the increasing sonority sequence \*(lb, \*bl). To give an explanation to those observations, Lowenstamm (1999) proposes that each word belonging to a major lexical categories must be preceded by a CV site. This site is subject to the same requirements of phonological government than any empty vocalic position. Departing from this suggestion, Lowenstamm shows that the initial site in French and in languages presenting the same consonantal initial configurations (type 1) has to be properly governed, while in Arabic for example where the initial clusters are free (type 2), the site is not always governed. What are the implications of such a theory?

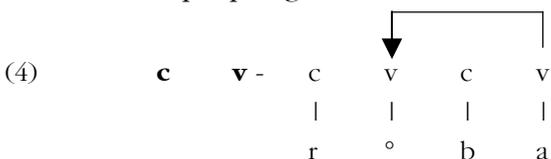
Since the initial site has to be properly governed in languages of type 1, we expect to find phonological arrangements allowing this proper government. This is what we observe in French for example. The consonantal word initials are limited to two configurations: a simple one (e.g. [fât] *fente* ‘crack’) or a complex one with increasing sonority (e.g. [bra] *bras* ‘arm’).

Lowenstamm (1999) adopts an idea put forward in Scheer (1996) suggesting that clusters formed by an occlusive and a liquid constitutes a closed domain. The result of the interaction between the two consonants is the possibility for the vowel captive by this relation to stay empty and to be transparent to proper government. The situations in (3a) and (3b) are then obtained.



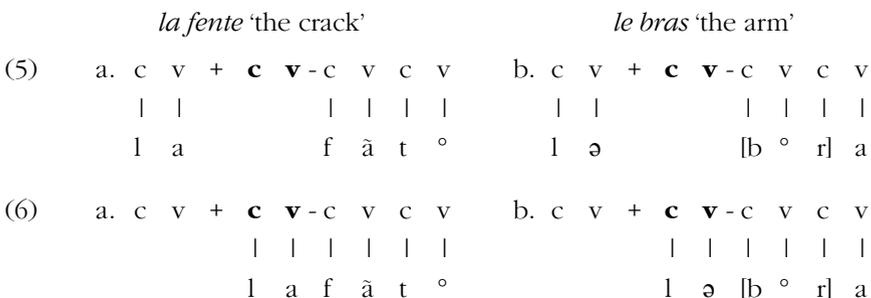


As the CV initial theory stipulates that French has a site always properly governed, both preceding situations are the only ones possible. A word like \*[rba] is ungrammatical in type 1 languages because it blocks the proper government of the initial site, as (4) shows.



The representation in (4) can only belong to languages of type 2, where the initial site may stay ungoverned. Also, both (3a) and (3b) can represent words from type 2 languages. It follows from the government of the initial CV that lexical material can become embedded in it. Lowenstamm (1999) gives the example of cliticization in French and Biblical Hebrew. We will present a brief summary of his argumentation here in order to show that there can be a clear manifestation of the initial site. For more details, the entire article should be consulted.

Remember that French is a type 1 language and has an initial CV site that is always governed. The definite articles in French are reputed to cliticize onto the nouns they determine. This cliticization is expected to take place in the CV initial site because it is always properly governed. (5) and (6) show this process.



We can see in (5) that the clitic comes with his own CV, and in (6) that it deserts its CV to join the initial site of the noun on which it cliticizes. The CV left behind in (6) plays the role of the new initial site and is ready to host other lexical material.

For the type 2 languages, the behaviour is different. Those languages can manage an ungoverned initial site by phonological operations. Lowenstamm presents the case of the definite articles in front of some plurals of the segholate class in Biblical Hebrew. When the article *ba* is prefixed to those nouns, the situation is similar to the one in (5) in the sense that the prefix has its own CV.

(7) *ba* + *klabim* 'the dogs'

|    |   |   |     |    |     |    |   |   |   |   |   |   |   |   |   |   |   |   |  |
|----|---|---|-----|----|-----|----|---|---|---|---|---|---|---|---|---|---|---|---|--|
| a. | c | v | + c | v  | - c | v  | c | v | c | v | c | v | c | v | c | v | c | v |  |
|    |   |   |     |    |     |    |   |   |   | \ | / |   |   |   |   |   |   |   |  |
|    | h | a |     | [k | °   | l] | a | b |   | i |   | m | ° |   |   |   |   |   |  |

*ba* + *rqahm* 'the spices'

|    |   |   |     |   |     |   |   |   |   |   |   |   |   |   |   |   |   |   |  |
|----|---|---|-----|---|-----|---|---|---|---|---|---|---|---|---|---|---|---|---|--|
| b. | c | v | + c | v | - c | v | c | v | c | v | c | v | c | v | c | v | c | v |  |
|    |   |   |     |   |     |   |   |   |   | \ | / |   |   |   |   |   |   |   |  |
|    | h | a |     | r | °   | q | a | ħ |   | i |   | m | ° |   |   |   |   |   |  |

(from (26), in Lowenstamm 1999)

The status of the site is different between (7a) and (7b). In (7a), the initial CV site is properly governed by the vowel /a/ of the base. As we said before, lexical material can incorporate the site if and only if it is properly governed. We then expect to find the cliticized form *\*baklaBiim* for (7a), where the *ba* simply becomes integrated into the initial site of the noun. But it is not the case. To understand what happens, it's advisable to look at the situation in (7b). The representation in (7b) shows that the initial site is ungoverned. The vowel /a/ of the base has to legitimize the empty vocalic position between the /r/ and the /q/ and therefore cannot legitimize the CV site. The result of this situation is the impossibility for the article to move into the site. We have then a case of an empty CV position that became part of the word. This CV cannot remain silent and have to undergo phonetic

interpretation. Two strategies are available: compensatory lengthening or gemination. In (7b), the vowel /a/ of the prefix lengthened on the vocalic position of the initial site. This is illustrated in (8b). The situation in (7a) can be resolved exactly like the one in (7b). Lowenstamm (1999) assumes the «Uniformity Convention», stipulating that *for any given language, cliticization operates in uniform fashion with respect to the licensing status of the host site* (p.12). So, Biblical Hebrew has cases of licensed initial site and cases of unlicensed initial site. The uniformity condition will make the CV site remains unlicensed throughout the language. What we observe for the cliticization in (7a) is the gemination of the first consonant /k/ of the noun to legitimize the CV site. This is illustrated in (8a).

(8) Gemination

a. c v + c v - c v c v c v c v c v  
 | |     ↖     | | | | | \ / | |  
 h a         k ə l a b     i     m °  
                   [hakkəlaβiim]

*Compensatory lengthening*

b. c v + c v - c v c v c v c v c v  
 | |     ↖     | | | | | \ / | |  
 h a         r ə q a ḥ     i     m °  
                   [haarəqahim]

(= (28a) et (28b), in Lowenstamm 1999)

It comes out of our discussion that the differences between languages, as for clusters at the beginning of words, are explainable by the presence of an initial site. Whereas we can observe a clear phonological manifestation of the initial CV site in languages of type 2 like Arabic and Biblical Hebrew, it is not the case for languages of type 1 like French and English. We can therefore question the presence of this CV site in languages where it is always properly governed and where no phonological operations of gemination or compensatory lengthening are expected. How can we verify the

presence or not of this site at the initial of French nouns? In the next section, we will put forth facts from French allowing us to answer this question.

## 2. THE DIMINUTIVE REDUPLICATION

Reduplication is a strategy of word formation used in most languages. This process imply that a part of a word or an entire word become affixed to itself to create another derived word. In French, we find cases of reduplication. There are two principal types: diminutives and onomatopoeias. We will leave the study of the latter for another research. For this paper, we will concentrate on diminutive (often figurative) formations by reduplication. Those nominal reduplications in French are characterized by the copy of the first consonant and the following vowel at the left of the base noun. Here are some examples in (9).

|     |           |                  |                                  |
|-----|-----------|------------------|----------------------------------|
| (9) | bébête    | 'little animal'  | from <i>bête</i> 'animal'        |
|     | bibiche   | 'little darling' | from <i>biche</i> 'doe, darling' |
|     | bobosse   | 'hunchback'      | from <i>bosse</i> 'hump'         |
|     | fifille   | 'little girl'    | from <i>fille</i> 'girl'         |
|     | gougoutte | 'little drop'    | from <i>goutte</i> 'drop'        |
|     | guéguerre | 'squabble'       | from <i>guerre</i> 'war'         |
|     | poupoule  | 'chick (woman)'  | from <i>poule</i> 'hen'          |
|     | susucré   | 'little sugar'   | from <i>sucre</i> 'sugar'        |

What are the exact functions assumed by this type of reduplication? First, we can attribute a morphological function to reduplication. Béchade (1992:157-158) notices that we can legitimately consider different processes, other than derivation or composition, put forward by French to enrich its vocabulary, as integrating a general system of word formation. The words in (9) are then clear cases of new word formation. Moreover, morphological processes, unlike phonological processes, often compete, as shown by Bauer (2001). This is what we observe in French with the minimal pairs *gougoutte*

‘little drop’ / *gouttelette* ‘droplet’ and *poupoule* ‘chick (woman)’ / *poulette* ‘lass’, the suffix *-ette* being in French the usual diminutive suffix. Each word of these pairs constitutes nominal derivation bringing the diminutive meaning to a common base.

Morphological derivation comes with a principle of semantics compositionality, stipulating that the meaning of the derived word should be the sum of the meaning of the parts. In other terms, we expect to find the meaning of the affix added to the meaning of the base in the semantics of the new word. When we look at nouns in (9), we see that this principle is respected. The noun *gougoutte* for example is the result of a derivation on the noun-based *goutte* ‘drop’. The semantics of the complex noun obtained is *petite goutte* ‘little drop’, that is to say a combination of the meaning of the base and of the reduplicated affix. We can therefore conclude that there is a morphological role to this type of reduplication in French.

Second, can we provide a phonological function to reduplication? The answer is linked to the theoretical frame we adopt. Remember that Lowenstamm (1999) proposes that words of all major categories are preceded by an initial CV site. If we assume that the noun *poule* ‘hen’ for example has the representation in (10), the suggestion is that the initial CV is where the diminutive reduplication takes place in French.

- (10)            **c**    **v-**    c    v    c    v  
                                          |    |    |    |  
                                          p    u    l    °                    [pul] *poule* ‘hen’

We propose following Lowenstamm (2002) that the phonological content of the word is copied at the left and that there is an association of the segments to the site from left to right. This process is illustrated in (11). It explains why we don’t have a form like \*[pulpul], with the entire reduplication of the noun. The initial site narrows reduplication to the first consonant and the following vowel only.

- (11)                                    **c**    **v-**    c    v    c    v  
                                                                  |    |    |    |  
                                                                  p    u    l    °                    [pupul] *poupoule* ‘chick (woman)’
-

Since the initial site is argued to be present in front of all nouns, the prediction is that this process is productive. Productivity is defined by Bauer (2001) and Bybee (1985, 1996) (among others) as the probability for a pattern to be used to form new words. Bybee shows that the amount of semantic and/or phonological restrictions is inversely proportional to the chances for a process to be productive. In the case of diminutive reduplications, we have seen that there was very few semantic restrictions, in the sense that the meaning of the reduplicate nouns is predictable on the meaning of the base. Moreover, all the nouns in (9) are formed by the phonological reduplication to the first CV of the base. We can notice here that there is only few phonological restrictions. Then, the process of reduplication presented in (11) has fulfilled all the conditions to be productive. The productivity of reduplication can be verified by the nouns in (12), all created on the model in (11).

|      |           |                             |
|------|-----------|-----------------------------|
| (12) | [bɔbɔt]   | from <i>botte</i> 'boot'    |
|      | [kakar̥t] | from <i>carte</i> 'card'    |
|      | [ʃɛʃɛz]   | from <i>chaise</i> 'chair'  |
|      | [fœfœ]    | from <i>feuille</i> 'sheet' |
|      | [fɛfɛs]   | from <i>fesse</i> 'buttock' |
|      | [papat]   | from <i>patte</i> 'leg'     |
|      | [pɔpɔrt]  | from <i>porte</i> 'door'    |
|      | [tɛtɛt]   | from <i>tête</i> 'head'     |

Of course, without a particular context, those words are hardly acceptable. Nevertheless, a French speaker will understand *petite chaise* 'little chair' if he hears [ʃɛʃɛz], as for *petite patte* 'little leg' if he hears [papat]. For example, the noun \*[patpat] for 'little leg', formed by the reduplication of the entire word could not be found. Since the semantic interpretation is predictable and since the reduplication is always restricted to the first open syllable in the creation of a new word, we can talk about a productive process. Productivity of the diminutive reduplication seems to be explained by the presence of an initial site.

Speaking in terms of a CV, the prediction that only the series Consonant-Vowel will be reduplicated is made. We have suggested in (11) that reduplication operates by an association from left to right of the copied phonological content in front of the base. Because the reduplication site has a restricted size, we predict that words beginning with a consonant cluster cannot reduplicate. This question will be the object of the next section.

### 3. THE CONSONANT CLUSTER CASE

So far, we have seen that nouns beginning with a simple consonant could be derived in diminutives by the reduplication to the left of the base of the first consonant and the following vowel. What we have to verify now is if reduplication is possible on nouns beginning with a consonant cluster and if this reduplication observes the same rules than in the preceding cases.

In search of an answer, we presented a list of nouns with initial consonant cluster to a group of French speakers and we asked them to form their diminutives by reduplication on the model illustrated in (11). The results of this query are resumed in (13). They are divided in two strategies of derivation. The first one in (13a) consists of a complete reduplication of the base. The second one in (13b) presents a reduplicant that is the base without the final consonant.

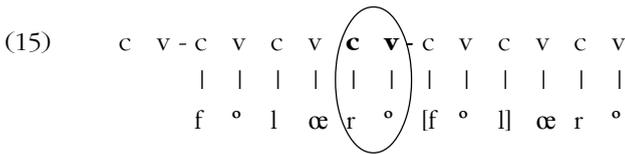
- |         |                        |             |    |                          |            |
|---------|------------------------|-------------|----|--------------------------|------------|
| (13) a. | <i>bras</i> 'arm' :    | bras-bras   | b. | <i>cloche</i> 'bell' :   | clo-cloche |
|         | <i>clé</i> 'key' :     | clé-clé     |    | <i>fleur</i> 'flower' :  | fleu-fleur |
|         | <i>fruit</i> 'fruit' : | fruit-fruit |    | <i>glace</i> 'ice' :     | gla-glance |
|         | <i>pluie</i> 'rain' :  | pluie-pluie |    | <i>place</i> 'place' :   | pla-place  |
|         | <i>pois</i> 'pea' :    | pois-pois   |    | <i>plume</i> 'feather' : | plu-plume  |

The diminutive formation illustrated by the first strategy meets another process existing in French. Indeed, we find a type of reduplication implying the complete copy of the base. Examples of this type of derivation are *train-train* 'routine' from *train* 'train' and *trou-trou* 'ornament with holes' from *trou* 'holes'. As these examples

show, words with a complex consonantal initial can be totally reduplicated. The difference with the ones in (13a) is the semantism. In (13a), the semantic interpretation is clearly diminutive, this is not the case for the two latter examples, *train-train* does not mean little train. Nouns of the kind of *train-train* are not very numerous in French. In fact, this type of reduplication is not productive at all and is rather lexicalized.

We have to wonder if the process described in (13a) is another kind of reduplication, on the model of *train-train* and then totally different than the reduplication described in the preceding section, or if it falsifies our hypothesis on the role of the initial site or even its presence. Two kinds of analysis can be provided for the data in (13a). First, we can say that the initial CV site has nothing to do with reduplications and that those diminutive formations are simply formed by the copy at the left of the first open syllable. This analysis would also explain why we have [papat] and not \*[patpat], for example. The problem is that in a theoretical framework like Government Phonology, consonants in a cluster are separated by an empty vocalic position. The implication here is that reduplications would be of variable size depending on the base. For illustrations, a word like *patte* ‘leg’ will reduplicate a CV, and a word like *bras* ‘arm’ will reduplicate a CVCV. We would prefer a process that operates in an uniform fashion. Second, we can say that reduplications in (13a) imply the reduplication of the entire word with its skeleton and the initial CV. We would then have an analysis similar to the one of *train-train*. A support for this second analysis resides in the results in (13b). If we try to illustrate in (14) and (15) what is going on with the reduplication of the noun *fleur* ‘flower’, we see that the initial site plays a role.

$$(14) \begin{bmatrix} c & v & -c & v & c & v & c & v \\ & | & | & | & | & | & | & | \\ \text{copy} & & [f & \text{ }^\circ & l] & \text{ }^\circ & r & \text{ }^\circ \end{bmatrix} + \begin{matrix} c & v & -c & v & c & v & c & v \\ & | & | & | & | & | & | & | \\ & [f & \text{ }^\circ & l] & \text{ }^\circ & r & \text{ }^\circ & \end{matrix}$$



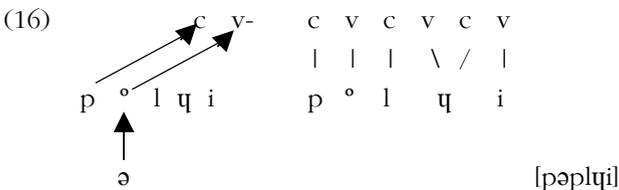
In (14), the noun *fleur* is reduplicated entirely. Since its initial CV site is properly governed, lexical material can be embedded in it. That is what happens in (15). The difference here is that the association is not from left to right and looks more like compounding. The copy merged within the base via the initial CV site. Government Phonology predicts the ungrammaticality of such a form. Scheer (2000) proposed that liquids have to be legitimized by a following realized vowel, except for the final position. In (15), the final liquid /r/ of the base is followed by an empty vocalic position but a final one. On the other hand, the consonant /r/ of the copy has now become a part of the word and has to be legitimized. This is not the case since the vocalic position following it is empty. Two actions can be envisaged to correct the situation: a schwa epenthesis in the vocalic position to legitimize the final /r/ of the copy or a deletion of the consonant /r/. The solution chosen is the most economical, the deletion of the final consonant of the copy.

It seems that the process involved in the diminutive formation in (13b) is the same than in (13a). This type of reduplication is not very productive in French as we said earlier. In fact, very few reduplications of entire words bring a diminutive semanticism. It is rather used to mark onomatopoeias (ex. *blabla* ‘smooth talk’, *froufrou* ‘rustle’). We will not discuss those cases more in this paper. What can be said about the reduplications described here is that it implies the entire copy of the base. At first sight, this strategy raises doubts about the hypothesis of an initial CV as the site of diminutive reduplication in French. Actually, on surface, we observe the reduplication of the first two consonants and the following vowel. We have presented in (14) and (15) what we believe to be the best analysis of those reduplications. The thing is, we still don’t have any morphological use of the initial site with nouns beginning with a consonantal cluster. The two next sections are concerned with this fact.

### 3.1. Diminutives in Canadian French

Data from Canadian French are particularly interesting for the study of diminutive formation by reduplication. The same nouns that served for the experience resumed in (13) have been presented to speakers of Canadian French, with the identical instructions, that is to form the diminutives of those nouns on the model (11). The results contrasts with the ones obtained for the French speakers. Lets see for example the noun *pluie* ‘rain’. Remember in (13a) that the diminutive of this noun, *pluie-pluie* ‘little rain (fig)’, was formed by the entire reduplication of the base. This is not what happens for some of the Canadian French speakers. Rather, the initial CV site is used.

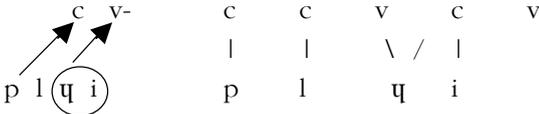
The noun *pluie* is preceded by an initial CV. The reduplication, as we describe it in section 2, consists of a copy of the phonological content at the left of the base and of an association from left to right of the segments on the segmental positions C and V of the site. The prediction that we made is that if the initial site is used to form diminutives, forms like \*[plɥiplɥi] *pluie-pluie* or \*[plyplɥi] *plu-pluie* or even \*[plɥiplɥi] *pli-pluie* are not expected. And this is the case for the majority of the Canadian French speakers questioned. Instead, [pəplɥi] is obtained, with the copy of the first consonant and a following schwa. So, we suggest that what is associated to the initial site corresponds to the first consonant /p/ and the empty vowel. The OCP (obligatory contour principle), preventing the two /p/’s being adjacent, does the rest and forces the schwa epenthesis into the empty vocalic position. This is illustrated in (16).



In fact, cases of reduplication on words with complex consonantal initial are extremely rare (at least, we haven’t found much). Here is a small list of our findings in (17).

- (17)            [bøbra]            from *bras* ‘arm’  
                   [føfløær]        from *fleur* ‘flower’  
                   [føfrũil]            from *fruit* ‘fruit’  
                   [gøglas]            from *glace* ‘ice’  
                   [pøplũil]            from *pluie* ‘rain’  
                   [pøplym]            from *plume* ‘feather’

The data in (17) put forward a fact discussed in the first section, that is if words of all languages are represented at a segmental level by a strict consecution of consonants (C) and vowels (V). The Canadian French data offer an answer. What is striking about the representation of the diminutive in (16) is the epenthesis of a schwa after the /p/. This would not be expected within a theoretical framework stipulating that the two consonants of increasing sonority form a complex onset. We would expect a situation where the other vocalic segments could come occupy the V position, for example the /ɥ/ or /i/, as (18) shows.

- (18)            

The glide /ɥ/ or the vowel /i/ could associate with the vocalic position. We would then expect to find [pyplũil] or even [piplũil], but this is not the case for all our Canadian French informers. So, if we suppose that the word *pluie* contains an empty vocalic position between the /p/ and the /l/ and that reduplication implies the copy from left to right of the segments into the initial site, the /p/ and the empty vowel will be reduplicated. The schwa epenthesis is a consequence motivated independently from the reduplication process. From now on, we can assume that French is a strictly CVCV language.

In section 2, we have discussed the necessary conditions for a word formation process to be productive. So, we have seen that productivity was related to the semantic and phonological restrictions imposed by the model. Diminutive reduplication on nouns beginning with a consonant cluster is not productive and this is not surprising

if we consider the situation described in (16). Of all the phonological content of the base, only the first consonant is (really) reduplicate. The semantic value conveyed by the segments is partly lost in the process of reduplication. Moreover, we notice that there is an additional phonological restriction in (17) comparing to reduplications in (9) and (12): the schwa epenthesis. Defining productivity as we did, it is clear that reduplications on nouns with complex consonantal initial are less expected. Nevertheless, the structure of those diminutives in French is the result of the presence of an initial CV site, at least in front of the nouns. In search for cross-linguistics data supporting our hypothesis concerning an initial site, Ancient Greek also present interesting facts to be considered in this paper and will be the subject of the next section.

### 3.2. Perfective formation in Ancient Greek

A type of reduplication similar to the one describe previously operates in Ancient Greek to form perfective. It consists of a reduplication in front of the base of the first consonant followed by the vowel /e/. Here are some examples in (19) drawn from Froli (2001: 47).

- |         |                       |                                          |
|---------|-----------------------|------------------------------------------|
| (19) a. | leluka                | from <i>luo</i> 'to loose'               |
|         | pepaideuka            | from <i>paideuo</i> 'to educate'         |
| b.      | gegrap <sup>h</sup> a | from <i>grap<sup>h</sup>o</i> 'to write' |
|         | beblap <sup>h</sup> a | from <i>blap<sup>h</sup>to</i> 'to harm' |

The data in (19a) are not surprising and look like those in (9). Lowenstamm (2002) proposes that the perfective formation in Ancient Greek is realized on the initial CV site, exactly as the French diminutives, with a vowel fixed by the perfective paradigm. The form *leluka* is then obtained instead of *\*luluka* as we would expect if the vowel was also reduplicated. For the reduplications in (19b), the situation is similar to the one in (17). If we suppose that the initial CV is the site for perfective reduplication, consonant clusters won't be able to reduplicate. This is what is observed, represented in (20).



The difference between this process and (16) resides in the quality of the reduplicated vowel. As the vowel in (16) is realized by the epenthesis of a schwa in the empty vocalic position of the initial site, the vowel /e/ in (20) is constant and is ruled by the paradigm of the perfective formation.

If we resume what we tried to show until now, we can say that the presence of an initial CV site is motivated morphologically and has phonological consequences on reduplication. The illustration of the diminutive reduplications in French has shown that words can be reduplicated following a specific phonological pattern, that is to say by the copy at the left of the base of the first consonant and vowel. This pattern ensues from the initial CV, where the segmental positions are restricted to a consonantal segment and a vocalic segment. In addition, the diminutives in Canadian French and the perfectives in Ancient Greek also support this hypothesis. In the next section, we will try to bring out a recurrent question in the literature: the status of the phonological content of the aspirated and nasalized segments.

#### 4. ASPIRATES AND NASALS

What we will do now is reduplicate words containing aspirated segments (Ancient Greek) and nasalized segments (French). Those phonological items have been dealt with ambiguity in the literature. Some authors treat them as unique segments occupying a single segmental position, while some others think that they are groups of segments occupying several skeletal positions. Diminutive reduplication offers a new way of analyzing aspirates and nasals, and thus clarifying this ambiguity, at least for Ancient Greek and French.

In Ancient Greek, there are verbs containing an aspirated consonant at initial. When the perfective reduplication is formed, the

data in (21) are obtained, with the remarkable impossibility to have the aspiration reduplicated.

- |      |                         |                                         |                                                         |
|------|-------------------------|-----------------------------------------|---------------------------------------------------------|
| (21) | kek <sup>h</sup> oreuka | * k <sup>h</sup> ek <sup>h</sup> oreuka | from <i>k<sup>h</sup>oreuo</i> ‘to dance’               |
|      | kek <sup>h</sup> areka  | * k <sup>h</sup> ek <sup>h</sup> areka  | from <i>k<sup>h</sup>airo</i> ‘to be delighted’         |
|      | tet <sup>h</sup> aumaka | * t <sup>h</sup> et <sup>h</sup> aumaka | from <i>t<sup>h</sup>aumad<sup>h</sup>o</i> ‘to admire’ |

The process of perfective formation by reduplication implies the copy of the first consonant and the insertion of the vowel /e/. If we analyze aspirates as only one phonological segment, we have to explain why the forms where the aspirated consonant is reduplicated are ungrammatical, by a rule of deletion of the aspiration, for example. Conversely, if we consider that the aspiration occupies its own segmental position, then no special rule is needed to explain why the occlusive only is reduplicated. These two stories are illustrated below.

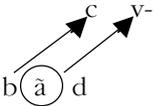
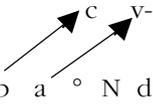
- (22) a.
- |  |                                         |               |    |                           |   |   |   |   |       |
|--|-----------------------------------------|---------------|----|---------------------------|---|---|---|---|-------|
|  |                                         | c             | v- | c                         | v | c | v | c | v     |
|  |                                         |               | e  |                           |   |   |   |   |       |
|  | (k <sup>h</sup> )o...                   |               |    | k <sup>h</sup>            | o | r | e | ° | u k a |
|  | /k <sup>h</sup> ek <sup>h</sup> oreuka/ | deletion rule |    | [kek <sup>h</sup> oreuka] |   |   |   |   |       |
- b.
- |  |                           |   |    |         |   |   |   |   |       |   |   |
|--|---------------------------|---|----|---------|---|---|---|---|-------|---|---|
|  |                           | c | v- | c       | v | c | v | c | v     | c | v |
|  |                           |   | e  |         |   |   |   |   |       |   |   |
|  | k ° h o...                |   |    | {k ° h} | o | r | e | ° | u k a |   |   |
|  | [kek <sup>h</sup> oreuka] |   |    |         |   |   |   |   |       |   |   |

In (22a), the only way to obtain the perfective form without the aspiration on the initial C is to add a rule, for example of deletion, complicating the process. A more economical suggestion is made in (22b), that is to consider the aspiration independent of the occlusive. We get back there the reduplication process as we described it in (20).

The same reasoning can be made concerning nasals reduplications in Canadian French. When a noun containing a nasal vowel is reduplicated to form the diminutive, the vowel is reduplicated without the nasality, as is shown by the examples in (23).

- |      |         |          |                            |
|------|---------|----------|----------------------------|
| (23) | [babād] | *[bābād] | from <i>bande</i> ‘group’  |
|      | [dedēd] | *[dēdēd] | from <i>dinde</i> ‘turkey’ |
|      | [ʒaʒâb] | *[ʒâʒâb] | from <i>jambe</i> ‘leg’    |

Many works on nasality, notably Paradis & Prunet (2000), discuss the number of segmental positions that nasal vowels must occupy. Data in (23) suggest that nasality is separate from the vowel it nasalizes, and so that nasals have to be considered as a composition of two segments. If not, forms like \*[bābād], \*[dēdēd] and \*[ʒâʒâb] would be expected. The two possible situations are illustrated in (24).

- |      |    |                                                                                   |   |    |   |    |   |   |
|------|----|-----------------------------------------------------------------------------------|---|----|---|----|---|---|
| (24) | a. |  | c | v  | c | v  |   |   |
|      |    |                                                                                   |   |    |   |    |   |   |
|      |    |                                                                                   | b | ã  | d | °  |   |   |
|      | b. |  | c | v  | c | v  | c | v |
|      |    |                                                                                   |   |    |   |    |   |   |
|      |    |                                                                                   | b | {a | ° | N} | d | ° |

When nasality is represented as part of the vowel (situation (24a)), the consequence is that the vowel AND nasality are reduplicated in one segmental position. When nasality is independent of the vowel (situation (24b)), then only the vowel can be reduplicated. The simplest way to explain why nasality is not reduplicated in the diminutive formation in French is by stipulating an analysis like the one in (24b).

More research needs to be done in the area of the segmental complexity in order to determine the true composition of aspirates and nasals. However, data from Ancient Greek and French examined above present interesting characteristics to be considered in this study. On one hand, they provide an additional argument for the presence of an initial site the size of a CV, and on the other hand, they show that it is better to give a bi-segmental status to the so-called complex segments like aspirated consonants and nasal vowels.

## 5. CONCLUSION

Many facts have been presented in this article to support the hypothesis of a CV site at the initial of words of all major categories. We have first seen that diminutive formation by reduplication in French is explainable by a morphophonological process occurring into the site. We have illustrated its productivity by showing that there were only very few semantic and phonological restrictions on this kind of reduplications. Moreover, the phonological pattern of the reduplication has been analyzed as being imputed to the initial CV site. We have then shown that the diminutives in French and the perfectives in Ancient Greek were formed by the strict reduplication of the first consonant and the following vowel of the base.

Predictions were made concerning the non reduplication of consonant clusters. A particular process existing in Canadian French was presented in order to confirm this hypothesis. In the case of nouns beginning with a complex consonantal initial, only the first consonant is reduplicated, followed by a schwa epenthesis. We have concluded that consonants of a cluster must be separated by an empty vocalic position. Thus, this is supporting the suggestion of Lowenstamm (1996) that words should be universally represented by a strict succession of consonants and vowels.

The hypothesis of an initial CV site provides two major openings. The first one is that it explains why consonant clusters cannot reduplicate. Indeed, we have seen in this paper that only the first consonant of a cluster is reduplicated, in Canadian French as in Ancient Greek. The second one concerns the composition of aspirates and nasals. The impossibility of reduplication in the case of an aspirated consonant and a nasal vowel suggested that those segments were occupying more than one segmental positions.

It is clear that this paper does not constitute the final word on this subject. However, it provides many leads on the treatment of the diminutive reduplication in French, as well as the perfective formation in Ancient Greek.

## FOOTNOTES

- <sup>1</sup> See also Bendjaballah (1995), Scheer (1996), Ségéral (1995).
- <sup>2</sup> The reader is referred to Kaye, Lowenstamm & Vergnaud (1985, 1990) for more information and precisions on Government Phonology.
- <sup>3</sup> This argumentation is based on the analysis of Czech vowel/ø alternations. Readers are asked to consult Scheer (1996) for demonstration of the complex situation in Czech.
- <sup>4</sup> This compensatory lengthening happens when the first consonant of the base resists gemination. (cf. Lowenstamm (1999:12))
- <sup>5</sup> We only suggest that a way of understanding the data in (13b) is to consider that the copy merged into the CV site, like we can imagine it does in compounding. Naturally, this proposition needs to be further explored, a thing we will not do here.
- <sup>6</sup> Some speakers of Canadian French produce a diminutive form where the first two consonants are reduplicated, followed by a schwa, for example [frəfrtʃi] for 'little fruit'. Those cases, not very representative though, won't be dealt with in this paper.
- <sup>7</sup> We have also run the test on French speakers, but since the nasals are less emphasize for them, we don't have a clear proof of the non reduplication of the nasalized vowel. Phonetic analysis should be made in order to have a precise answer.

## REFERENCES

- BAUER, L. *Morphological productivity*. Cambridge: Cambridge University Press. 2001. (Cambridge Studies in Linguistics, n. 95)
- BÉCHADE, H.-D. *Phonétique et morphologie du français moderne et contemporain*. Paris: Presses Universitaires de France, 1992.
- BENDJABALLAH, S. *Aspects du système verbal du berbère (kabyle)*. 1995. Doctoral (Dissertation) - Université Paris 7, Paris, France.
- BYBEE, J. L. Productivity, regularity and fusion: how language use affects the lexicon. In: SINGH, R. (Éd.). *Trubetzkoy Orphans: Proceedings of the Montreal Roundtable 'Morphology: Contemporary Responses'*. Amsterdam/Philadelphia: John Benjamins. 1996. p.247-269.
- BYBEE, J. L. *Morphology: A Study of the Relations between Meaning and Form*. Amsterdam/Philadelphie: John Benjamins, 1985.

FROLI, D. S. *De la lénition des codas initiales en grec. Étude diachronique de la spirantisation grecque dans un cadre CVCV*. 2001. DEA dissertation - Université de Nice, Nice, France.

KAYE, J.; LOWENSTAMM, J.; VERGNAUD, J.-R. Constituent structure and government in phonology. *Phonology* 7, n. 2, p. 193-231. 1990.

KAYE, J.; LOWENSTAMM, J.; VERGNAUD, J.-R. The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook* 2, p. 305-328, 1985.

LOWENSTAMM, J. Paramétrisation du CV initial (le bébé et l'eau du bain), paper presented to the *4èmes Journées Internationales du GDR CNRS1954 'Phonologies'*, Université Stendhal, Grenoble, France. 2002.

LOWENSTAMM, J. The beginning of the word. In: RENISON, J.R.; KÜHNAMMER, K. (Ed.). *Phonologica 1996: Syllables!?*, The Hague: Thesus, 1999. p. 153-166.

LOWENSTAMM, J. CV as the only syllable type. In: DURAND, J.; LAKS, B. (Ed.). *Current Trends in Phonology: Models and Methods*. Salford: European Studies Research Institute, 1996. p. 419-443.

PARADIS, C.; PRUNET, J.-F. Nasal Vowels as Two Segments: Evidence from Borrowings. *Language*, v. 2, n. 76, p. 324-357, 2000.

SCHEER, T. *De la localité, de la morphologie et de la phonologie en phonologie*. 2000. HDR dissertation - Université de Nice, Nice, France.

SCHEER, T. *Une théorie de l'interaction directe entre consonnes*. 1996. Doctoral dissertation - Université Paris 7, Paris, France.

SÉGÉRAL, P. *Une théorie généralisée de l'apophonie*. 1995. Doctoral dissertation - Université Paris 7, Paris, France.

## AUTHOR NOTE

I am grateful to Jean Lowenstamm for discussions on the differences existing between France French and Canadian French and on theoretical issues dealt in this paper.

This research was supported by a doctoral fellowship accorded by the *Fonds de recherche sur la société et la culture* from the government of Quebec.

# **Breve estudo da estrutura silábica do português arcaico: a posição do glide\***

Fernanda Elias Zucarelli  
Faculdades Integradas FAFIBE

## **Abstract**

The aim of this study of Medieval Portuguese vowel sequences is the study of vowel sequences in the Cancioneiros and other Medieval Portuguese documents. The phonological status of these vowels is defined as hiatus or diphthongs. All data have been submitted to auto-segmental and metrical analyses, with particular reference to the general phonological structure of Medieval Portuguese, according to the literature.

In the metrical approach, only the vocalic element can occupy the Nucleus position.

In falling diphthongs, all types of glides are analysed as Coda elements. Finally, it is necessary to note that MP has a significative amount of ambissyllabic sequences of vowels.

## 0. INTRODUÇÃO

O próprio título deste trabalho deixa claro o seu objetivo: trata-se de um breve estudo acerca da estrutura silábica do português arcaico (de agora em diante PA), no seu período trovadoresco, enfocando, principalmente, a posição do glide. Para tal, foram focalizadas todas as sílabas que ocorrem no *corpus* selecionado. Este foi constituído de 100 (cem) cantigas de amigo e de amor, extraídas do *Cancioneiro da Biblioteca Nacional de Lisboa* (de agora em diante CBN) e do *Cancioneiro da Ajuda* (de agora em diante CA). A análise foi baseada nas teorias fonológicas não-lineares, com o intuito de sugerir uma planilha silábica para o PA e estudar a posição do glide nessa planilha.

## 1. METODOLOGIA

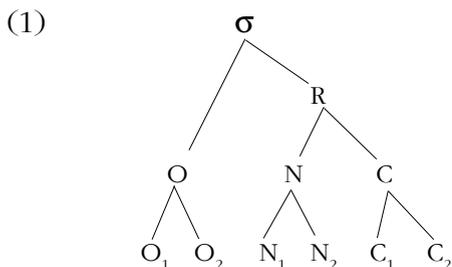
Como objeto de análise, foram consideradas 100 cantigas de amigo e de amor extraídas, respectivamente, do CBN e do CA. As cantigas que constituem o *corpus* foram decifradas cuidadosamente com o respaldo constante de edições críticas já publicadas desses textos (NUNES, 1973; PIMPÃO, 1942; SPINA, 1991; GONÇALVES & RAMOS, 1985). Inclusive, tomou-se o cuidado de desenvolver as abreviaturas que aparecem com freqüência nos Cancioneiros sob análise.

Posteriormente, as cantigas foram escandidas e foi feito um minucioso levantamento das estruturas silábicas e dos encontros vocálicos que aparecem no *corpus*. Quando se tem como objetivo a investigação de elementos de natureza não meramente segmental – como no caso a descrição de estruturas silábicas – de um período de uma língua quando ainda não havia tecnologia suficiente para o arquivamento e a transmissão de dados orais, a possibilidade de escolha entre material poético e não-poético para constituição do *corpus* não é viável. Como os textos remanescentes do PA são todos registrados

em um sistema de escrita de base alfabética, sem notações especiais para os fenômenos prosódicos (a não ser, na prosa, certos tipos de segmentação, na escrita, representados pela inclusão de espaços em branco entre as palavras, troca de linha, etc., que podem ajudar um pouco a clarificar essa questão), fica praticamente impossível de serem extraídas informações como as que estão sendo aqui buscadas, a respeito da estruturação silábica do português deste período, a partir de textos escritos em prosa. Como se pode ver, torna-se essencial o uso do material poético.

## 2. EMBASAMENTO TEÓRICO

Para o desenvolvimento deste trabalho, a estruturação das sílabas será comumente representada através de planilhas silábicas e pode ser apresentada em forma de árvore, como as sugeridas por Selkirk (1980) e Hogg & McCully (1987), o modelo segue abaixo em (1):



Assim, tentaremos, à luz da teoria métrica, partindo da análise das cantigas que compõem o nosso *corpus*, sugerir uma planilha silábica para o PA, indicando a posição do glide.

## 3. INTERPRETAÇÃO FONOLÓGICA DO GLIDE E A SUGESTÃO DE UMA PLANILHA SILÁBICA PARA O PA

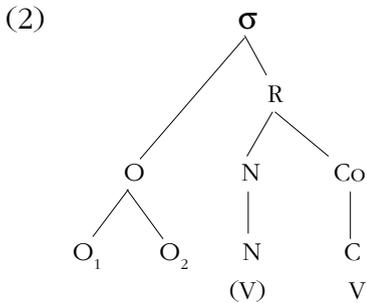
A hipótese de considerar planilhas silábicas com duas posições disponíveis no núcleo para vogais e glides, na verdade, auxilia na resolução de problemas, como a falta de lugar para “ancoragem” de vogais ou consoantes, mas, no PA, não existem problemas similares

a esse; porém, tal situação é bastante comum para outras línguas, como o inglês, por exemplo, que admite seqüências como *graind*, estudadas por Hogg & McCully (1987: 41).

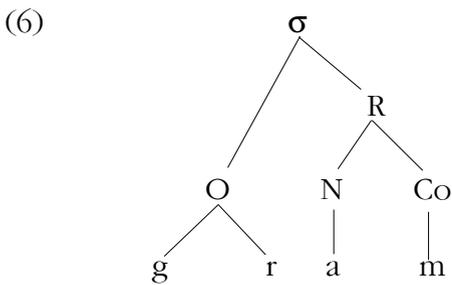
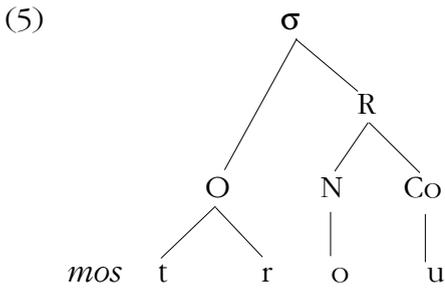
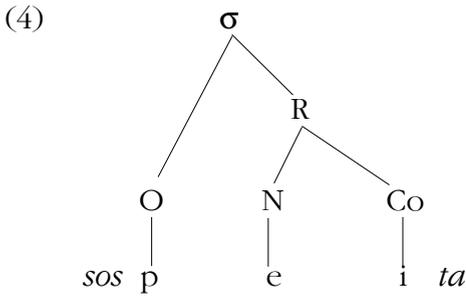
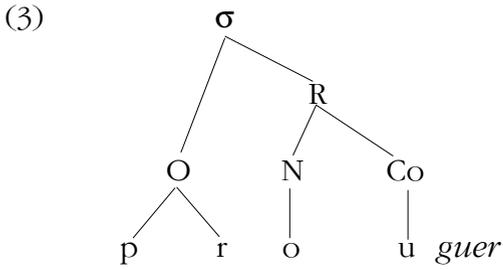
Segundo os dados extraídos do *corpus*, o PA apresenta no máximo estruturas silábicas com 4 constituintes, como se pode conferir em Bernardineli (1999: 393) e não apresenta longas seqüências de vogais e consoantes como as permitidas pela planilha acima que prevê sílabas com até 6 constituintes; nesse sentido, essa hipótese deve ser desconsiderada.

Assim, para o PA, uma segunda hipótese, a de que o glide estaria na coda, é a mais adequada, pois daria conta das estruturas silábicas da língua em questão e não deixaria posições ociosas na planilha; por isso, torna-se necessário apresentar uma planilha prevendo algumas restrições, para que não haja possibilidade de organizar estruturas silábicas agramaticais para o PA.

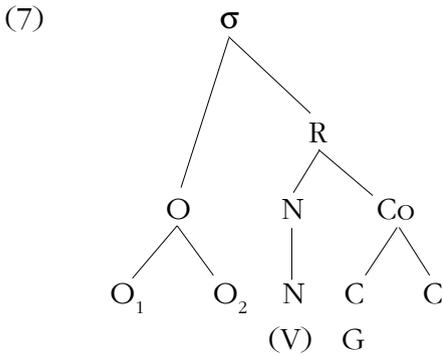
Lopez (1979) sugere, para o português brasileiro, um molde silábico que pode ser representado através da planilha silábica (2), como mostra Collichonn (1996: 113):



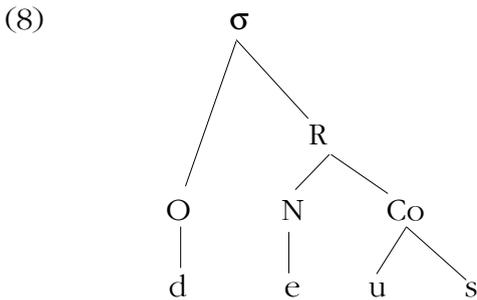
Esse molde é bastante reduzido em relação aos outros sugeridos, mas é capaz de representar quase todas as sílabas máximas (isto é, as sílabas possíveis com o maior número de segmentos) do PA, como mostram (3), (4), (5) e (6):



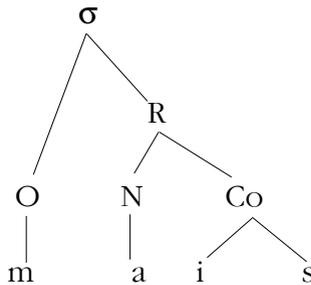
Ficam sem explicação, a partir da planilha considerada, somente ocorrências do tipo *deus*, *mais*, *meus*, *pois*, já que nesses casos temos um ditongo decrescente acrescido de um *s*, o qual ficaria sem lugar de ancoragem na estrutura silábica apresentada para o PA, com base em Lopez (1979). Surge, portanto, a necessidade de considerar as duas posições disponíveis na coda sugeridas pela estrutura apresentada em (1). Logo, podemos concluir que, para o PA, a planilha silábica<sup>1</sup> poderia ser apresentada com os seguintes constituintes:



Assim, temos uma estrutura que representa as seqüências vocálicas de maneira adequada, demonstrando que o núcleo da sílaba, sempre uma vogal, tem *status* diferenciado perante os outros constituintes, inclusive, perante os glides, que ficariam sempre na coda, como é possível ver em (8) e (9).

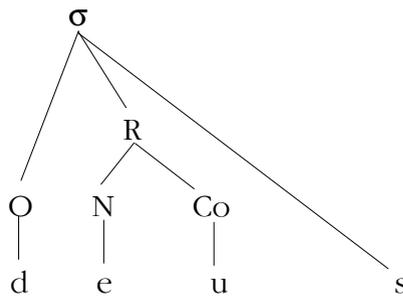


(9)

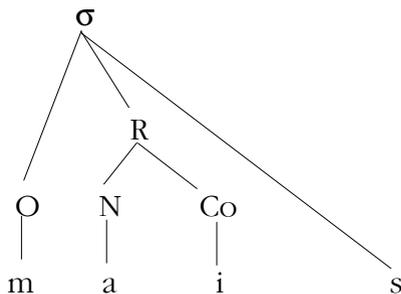


Outra hipótese é considerar o /s/ final extra-silábico, como fizeram Fikkert & Freitas (1997), (*apud* SANTOS, 2001, p. 81). Neste caso, as planilhas (8') e (9') poderiam ser reescritas como:

(8')



(9')



Bernardineli (1999, p. 393), após detalhado estudo das estruturas silábicas presentes em uma cantiga de D. Dinis, apresenta um quadro com os principais tipos silábicos, número de ocorrências e exemplos. Para verificar se a planilha proposta em (44) dá conta de todas as ocorrências do PA, representaremos nas planilhas abaixo alguns exemplos apresentados por Bernardineli.<sup>2</sup>

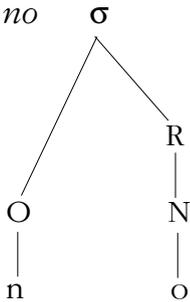
(10)

Tipo de sílaba: V

ejemplo: *e*

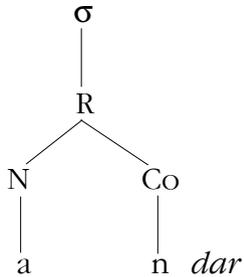
(11)

Tipo de sílaba: CV

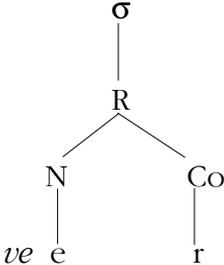
ejemplo: *no*

(12)

Tipo de sílaba: VC

ejemplo: *andar* e *veer*

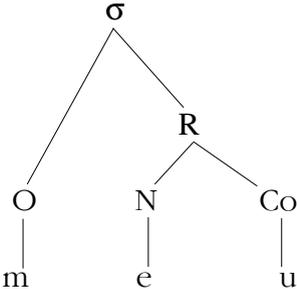
(12')



(13)

Tipo de sílaba: CVV

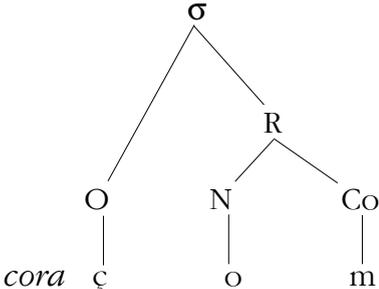
exemplo: *meu*



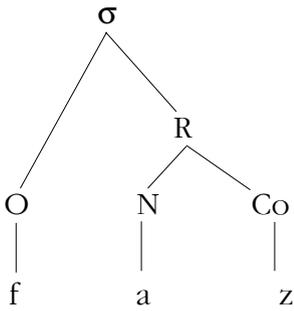
(14)

Tipo de sílaba: CVV

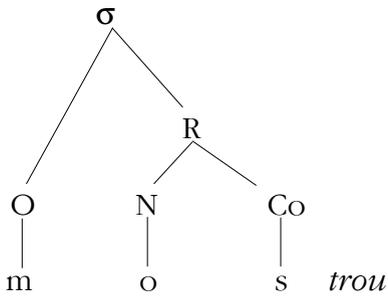
exemplos: *coraçom*, *mostrou* e *faz*



(14')



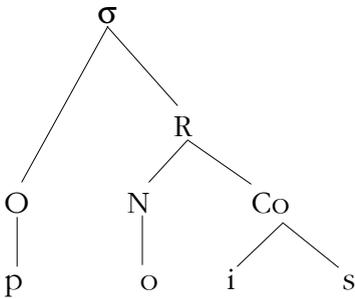
(14'')



(15)

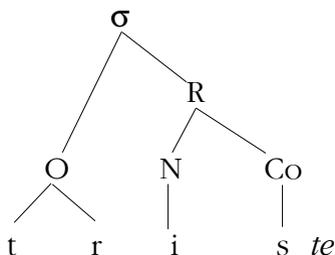
Tipo de sílaba: CVVC

exemplo: *pois*



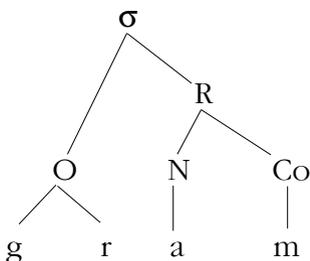
(16)

Tipo de sílaba: CCVC

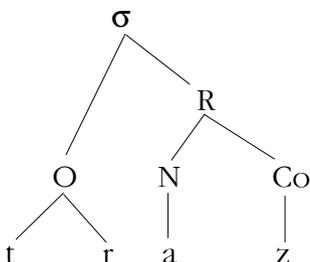
exemplo: *triste*

(17)

Tipo de sílaba: CCVC

exemplos: *gram* e *traz*

(17')



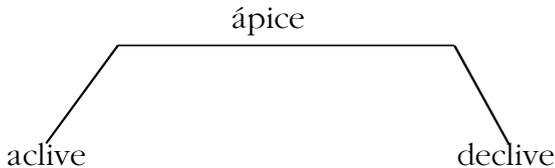
Temos, como se pode ver, os elementos /y/ e /w/ *trabalhando* como consoantes, travando sílabas, como em *par*, *pai*, tornando tais sílabas pesadas, proporcionando, inclusive, a atração do acento,<sup>3</sup> no caso do PA, como explica Massini-Cagliari (1999).<sup>4</sup> Nessa perspectiva,

devemos interpretar o glide com o mesmo *status* das consoantes que proporcionam o travamento de sílabas, ou seja, posicionado na *coda* da estrutura silábica. Além disso, Bisol (1994), ao discutir os ditongos derivados, sugere a formação de ditongos (diacronicamente falando) a partir de consoantes geminadas, defendendo que o glide deriva da consoante que ocupava a coda, travando a sílaba.

Câmara Jr. (1969: 26-27), ao esclarecer sobre a grande predominância de sílabas livres ou abertas em português, afirma:

Sílabas travadas ou fechadas são muito menos freqüentes e com limitação muito grande das consoantes que podem figurar no aclave, isto é, como decrescentes. Em compensação, as vogais /i/ e /u/ podem figurar nesta parte da sílaba como decrescentes e assilábicas (em transcrição fonológica, respectivamente, /y/ e /w/). Constituem a vogal silábica o chamado ditongo decrescente, como em *pai* e *pau*, que se opõem distintivamente a *pá*.

Ao analisar as sílabas dos ditongos como travadas, considerando *V/y,w/* como uma modalidade de sílaba travada, assim como *V/z/*, *V/r/* ou *V/l/*, Câmara Jr. (1969) admite uma análise VC, o que implicaria considerar o glide na coda. Além disso, o estudioso prevê a sílaba sendo constituída por um *active*, um *ápice* e um *declive*. Collischonn (1996: 111) sugere o seguinte gráfico:



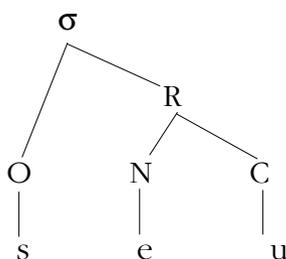
e o interpreta:

O *ápice* é constituído por uma vogal. O *active* é constituído por uma ou duas consoantes. O *declive* é constituído por uma das seguintes consoantes /S/, /r/, /l/ ou pela semivogal /y,w/. Além destas, considera a consoante nasal no declive, já que interpreta as vogais nasais como sendo fonologicamente “vogal fechada por consoante nasal”.

Tal interpretação de Collichonn (1996) nos permite acreditar que a autora defende a posição do glide na coda e sugere que Câmara Jr. (1969) também defendia essa hipótese,<sup>5</sup> que pode, realmente, ser sustentada e sugerida como sendo a mais viável também para o PA.

Cagliari (1998, p.49) afirma que, no português brasileiro, *a ditongação ocorre por uma ramificação da Rima, com a vogal ocupando a posição de Núcleo e a semivogal a de Coda*. Além disso, tal interpretação pode ser reforçada pela teoria desenvolvida por Spencer (1996) que sugere a intuição do falante acerca dos constituintes silábicos e apresenta a vogal como elemento central da sílaba, ou seja, a vogal desempenhando papel de alicerce para a sílaba, tratando-se assim, do elemento mais proeminente. Sendo o glide uma semivogal, pode-se concluir que terá um papel secundário, com menor proeminência; para que o falante possa distinguir um ditongo de um hiato, o glide deverá ocupar uma posição na coda, como defendem os estudiosos acima citados. Como esta argumentação é válida também para o PA, conseqüentemente, teremos, na verdade, planilhas com apenas uma posição disponível no núcleo, como (18):

(18)



No PA, deparamo-nos com ocorrências como *meu, seu, sandeu*, representadas acima, que são idênticas aos ditongos atuais, mas a interpretação fonológica não é tão tranqüila quanto parece, visto que a qualidade desses encontros vocálicos pode ser discutida – o que tem importância crucial para a análise. Bisol (1989) distingue dois tipos de ditongos: os pesados (ditongos “verdadeiros”, irredutíveis – ex.: *reitor*) e os leves (redutíveis – ex.: *caixa*). Se, na estrutura silábica, ambas as vogais forem moraicais,<sup>6</sup> tratar-se-á de um ditongo pesado,

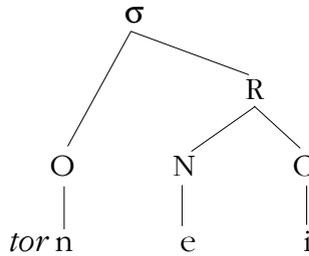
segundo Bisol. Ao contrário, se ambas as vogais estiverem subordinadas a uma mesma mora, então, a estrutura corresponderá a um ditongo leve.

Bisol (1994, p.123) retoma a análise feita em 1989 sobre os ditongos fonológicos (“verdadeiros”) e os ditongos fonéticos (“falsos”) e afirma:

O primeiro, invariante (*reitor, pauta*) está representado na estrutura subjacente por duas vogais, como um autêntico ditongo, enquanto o segundo, que ora se manifesta ora não (*peixe – pexe, feira – ferã*), possui na estrutura subjacente, apenas uma vogal, formando-se o glide em nível mais próximo à superfície.

Assim, temos nos verdadeiros ditongos uma rima ramificada, constituída de duas vogais, sendo que *a vogal alta dessa configuração manifesta-se foneticamente como glide*, como defende Bisol (1994, p.126). Tais ocorrências podem ser exemplificadas por árvores como (19):

(19)



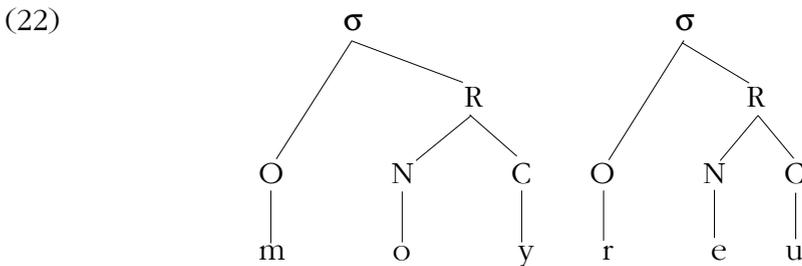
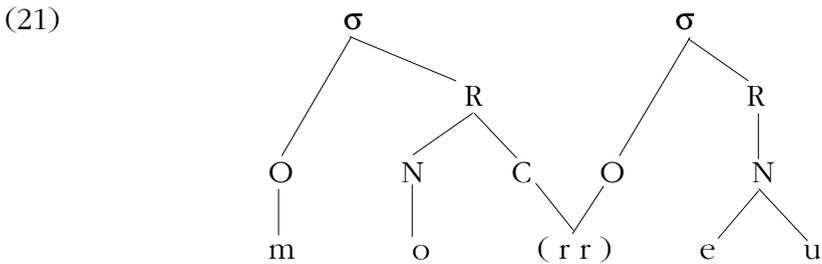
Já os “falsos ditongos”, segundo Bisol (1994, p.138), não existem na forma subjacente ou de base, mas são consequência de um processo de assimilação, portanto, podem ou não se realizar, sendo dependentes da criação de um glide epentético.

A diferença, portanto, entre os ditongos “verdadeiros” e os “falsos” está na forma de base, ou seja, são encontros vocálicos na estrutura superficial, mas não necessariamente na estrutura profunda; em outras palavras, os primeiros apresentam duas vogais na rima desde a estrutura subjacente, os segundos, não.

Existem ainda ocorrências, como a variação *moyreu* e *morreu*, muito recorrente e que aparece, por exemplo, nos versos 5 e 6 da cantiga 35 do CA, de autoria de Paio Soares de Taveirós:

- (20) Ay | mia | ſen | nor | a | ſſi | **mo**ȳ | **reu**  
 Co | mo | **mo** | **rreu** | quen | fōy | a | mar

Neste caso, como mostra Somenzari (2001), a seqüência “RR”, em *morreu*, é uma consoante geminada, ou seja, ocupa uma posição na *coda* da sílaba anterior e uma no *onset* da posterior; assim, pode-se dizer que, na ocorrência *moyreu*, o glide se forma da mora deixada pela consoante que trava a primeira sílaba (como se trata de uma consoante geminada, ela ocupa a *coda* da primeira sílaba e o *onset* da segunda ao mesmo tempo), trata-se, desta forma, na terminologia de Bisol (1994), de um ditongo derivado, cujo processo de formação está representado nas planilhas (21) e (22) :



#### 4. CONCLUSÃO

Assim, pode-se concluir que o glide deve ocupar posição na coda (pois as semivogais desempenham as mesmas funções das consoantes que travam sílabas, logo devem ocupar posição na coda). Além disso, foi possível sugerir uma planilha silábica para o PA –

reduzida, se comparada às planilhas sugeridas por Selkirk (1980) e Hogg & McCully (1987), pois sugere uma alteração: um único lugar de ancoragem no núcleo – que viabiliza a interpretação de todas as estruturas silábicas e não abre a possibilidade de serem geradas estruturas agramaticais para o PA.

## NOTAS

\* Esta comunicação é parte da Dissertação de Mestrado *Ditongos e Hiatos nas cantigas galego-portuguesas*, sendo uma das pesquisas que compõem o Projeto temático *Fonologia do Português Arcaico*, na Faculdade de Ciências e Letras (FCL) da UNESP, Campus de Araraquara, sob coordenação da Profa. Dra. Gladis Massini-Cagliari - financiado pela FAPESP (processo nº. 1997/12447-5).

<sup>1</sup> Onde V = vogal, C = consoante e G = glide.

<sup>2</sup> Não trabalharemos todos os exemplos de Bernardineli (1999), porque algumas ocorrências precisam de maiores discussões, como *lb'eu*, que é discutida no item 1.2.3. deste capítulo.

<sup>3</sup> Massini-Cagliari (1999, p.171) apresenta exemplos.

<sup>4</sup> Ver Massini-Cagliari (1999, p.89), que esclarece sobre possíveis divergências acerca de sílabas leves e pesadas entre diferentes línguas.

<sup>5</sup> Ver também Schane (1995).

<sup>6</sup> De acordo com o modelo métrico de Hayes (1995), para estabelecer o peso da sílaba é preciso contar o número de elementos no núcleo ou na rima, através da contagem de moras ( $\mu$ ), assim, a sílaba associada a uma mora é tida como leve e a sílaba associada a duas moras, ou mais, como pesada.

## REFERÊNCIAS BIBLIOGRÁFICAS

BERNARDINELI, A. Análise da estrutura silábica em uma cantiga de D. Dinis. *Estudos Lingüísticos*, São Paulo: GEL, v. 28, p. 391-397, 1999.

BISOL, L. Ditongos derivados. *D.E.L.T.A.*, v. 10, nº especial, p. 123-140, 1994.

CÂMARA JR., J. M. *Problemas de Lingüística Descritiva*. 2. ed. Petrópolis: Vozes, 1969.

CANCIONEIRO da Ajuda – Lisboa: Edições Távola Redonda. Edição Fac-similada. Cancioneiro da Ajuda. Lisboa: Sá da Costa, 1994.

CANCIONEIRO da Biblioteca Nacional (Colocci-Brancuti) Cod. 10991. Reprodução fac-similada. Lisboa: Biblioteca Nacional/Imprensa Nacional – Casa da Moeda, 1982.

COLLISCHONN, G. A sílaba em português. In: BISOL, L. *Introdução a estudos de fonologia do português*. Porto Alegre: Edipucrs, 1996.

GONÇALVES, E.; RAMOS, M. A. *A lírica profana galego-portuguesa*. 2. ed. Lisboa: Editorial Comunicação, 1985.

HOGG, R.; McCULLY, C. B. *Metrical phonology: a coursebook*. Cambridge: Cambridge University Press, 1987.

LOPEZ, B. S. *The sound pattern of Brazilian Portuguese (Cariocan dialect)*. Los Angeles: University of California, Ann Harbor, University Microfilms International. Tese (Doutorado), University of California, 1979.

MASSINI-CAGLIARI, G. *Do poético ao Lingüístico no ritmo dos trovadores: três momentos da história do acento*. Araraquara: Laboratório Editorial FCL/UNESP; São Paulo: Cultura Acadêmica, 1999.

NUNES, J. J. *Cantigas d'amigo dos trovadores galego-portugueses*. Lisboa: Centro do Livro Brasileiro. 1ª edição: 1926/1929, 1073

SELKIRK, E. O. *On prosodic structure and its relation to syntactic structure*. Indiana: IULC, 1980.

SOMENZARI, T. Estudo do *status* fonológico das consoantes duplas em Português Arcaico. Comunicação apresentada no II EDiP – II Encontro de Estudos Diacrônicos do Português. Araraquara: FCL, UNESP, 29 a 31 de agosto de 2001.

# **Traço fonológico: variação sob a perspectiva da multilinearidade**

**Teresinha de Moraes Brenner  
Universidade Federal de Santa Catarina**

## **Abstract**

The concept of natural class delineated in a geometric figure dates back to the Prague School. The idea was incorporated by multilinear phonology. Inserted on phonological different levels and plurilinear lines, dominated by specific categorial nodes, the phonological features are organized hierarchically. Particular contexts allow for the occurrence of phonological processes characterizing linguistic variation.

In the Portuguese language of Brazil, certain phenomena are explained by the substitution of features registered in related subgroups belonging to a broader category. Other processes allow for the substitution of elements by others in the hierarchy, relating to the degree of aperture and sonority – phonemic reinforcement or weakening. Finally, assimilation and dissimilation are illustrated and can be explained by segmental transparency or opacity. As a consequence, the geometry of the features re-covers the linguistic variation.

## 1. PROPOSIÇÃO

O presente artigo tem como meta a tentativa de explicar alguns fenômenos da variação do português do Brasil fundamentados em princípios fonológicos básicos da geometria dos traços. Numa breve retrospectiva, remonta-se a Sergeevič Trubetzkoy que introduz o conceito de classe natural, através de elementos aparentados, e os distribui em figuras geométricas. Retoma-se Noam Chomsky e Morris Halle que postulam uma única classe para consoantes e vogais baseados no critério articulatório. O trabalho pioneiro de George Clements, 1985, formula princípios da geometria dos traços conectados à fonologia multilinear.

A fonologia multilinear prevê a distribuição dos traços em categorias hierárquicas dispostas binariamente em planos opostos. As classes aparentadas podem se inserir no mesmo nível ou em níveis diferentes, determinando processos fonológicos específicos. Essa configuração faculta explicar alguns fenômenos do português do Brasil como o deslocamento do ponto de articulação no mesmo nível e em outros níveis. O enfraquecimento ou o reforço fonêmico justificam o último caso. Processos de assimilação e dissimilação implicam a opacidade e a transparência fonêmicas.

## 2. BREVE RETROSPECTIVA

Trubetzkoy, representante da Escola de Praga, no início do século XX, estipula um sistema de correlações, baseado nas oposições bilaterais proporcionais privativas neutralizáveis. Configura feixes de correlações da mesma classe. Em *Principes*, formula os sistemas do sânscrito e do grego antigo, delineando, respectivamente, feixes de quatro e três membros. A primeira língua registra duas correlações



Alfabeto Fonético Internacional que separa as categorias consoante e vogal no processo classificatório.

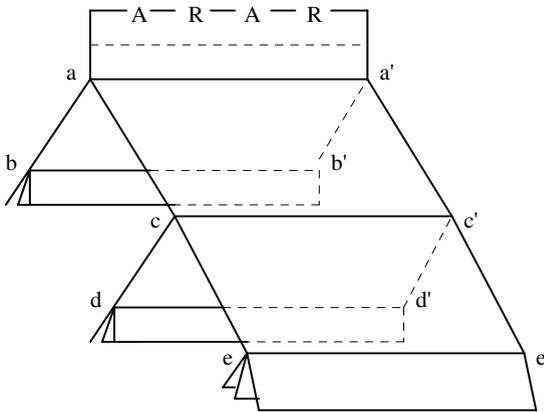
George Clements (1985) em *The geometry of phonological features*, em consonância com outros fonólogos da época, organiza o sistema de traços fonológicos, segundo os princípios da fonologia multilinear. Exclui o conceito de fonema como *bundle of features*. Inspirado na fonologia métrica, delimita a geometria dos traços num quadro de nós categoriais distribuídos em níveis hierárquicos. Os traços fonológicos constituem os elementos terminais de categorias dispostas binariamente em planos opostos em cada nível. Os processos fonológicos se realizam dentro dos limites de um mesmo nível ou pela articulação de níveis. O nó categorial mais alto, a raiz, se prende ao esqueleto de uma configuração multilinear. Assinala-se, portanto, uma independência relativa de cada nível, sendo os elementos mais incrustados os mais livres.

Os parâmetros articulatorios mostram um alto grau de independência entre si:

- (2) (a) configuração laringal;
- (b) grau de constrição da cavidade nasal (aberto/fechado);
- (c) grau e tipo de constrição da cavidade oral;
- (d) articulador ativo e articulador passivo.

Veja-se uma configuração multiplanária, cujos termos estão traduzidos para o português, traçada conforme a sugerida por Clements (p. 229), no artigo em questão:

## (3) Representação da geometria dos traços



aa' = raiz

cc' = nível supralaringal

ee' = ponto

bb' = nível laringal

dd' = modo

---- = esqueleto

Vários modelos paralelos têm sido propostos para explicar a geometria dos traços.

### 3. VARIAÇÃO: FENOMENOS DO PORTUGUES DO BRASIL

Inspirando-se no conceito de *classe natural*, formulado em Praga e desenvolvido na atualidade por fonólogos como Clements, procura-se explicar, através da geometria dos traços, o processo variacional do português do Brasil.

Questiona-se, inicialmente, a variante livre [ S ], em posição de *coda*, realizada como sibilante ou chiante. Sejam:

(4) (a) [mes] [meʃ] (mês)

(b) [meʒes] [meʒɐʃ] (mesas)

Constata-se que o processo variacional da *coda* se efetiva no mesmo nível: ponto de articulação. O falante, por opção ou por

confinamento cultural, pode selecionar no léxico, como elemento virtual, a realização [alveolar] ou [palatal]. Significa que as variantes de [S], [s] ou [ʃ], se encontram, quanto ao ponto de articulação, em planos distintos: [anterior] x [posterior]. Explicitando melhor, pode-se afirmar que a [coronal + ant] contrapõe-se a [coronal -ant]. Não se questiona aqui o enquadramento da [palatal] como [coronal] ou [dorsal] ou, ainda, como segmento complexo. Importa a oposição- anterioridade/ posterioridade, que permanece no sistema classificatório. Sejam:

(5) ponto de articulação da *coda fricativa*

Ponto de articulação

[s]  
|  
[+cor, + ant]

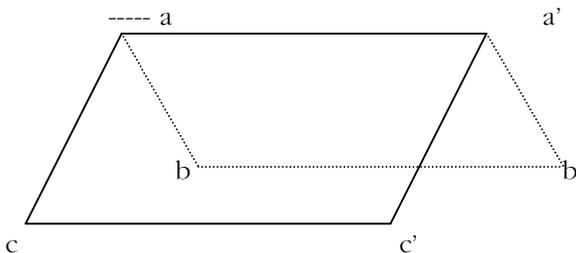
OU

Ponto de articulação

[ʃ]  
|  
[+cor, -ant]

A figura (6) esquematiza a bipartição do ponto de articulação em dois planos distintos, situando o processo no mesmo nível numa representação multiplanária. Confirme-se (conforme CLEMENTS, 1993, p. 134):

(6) distribuição do ponto de articulação em planos



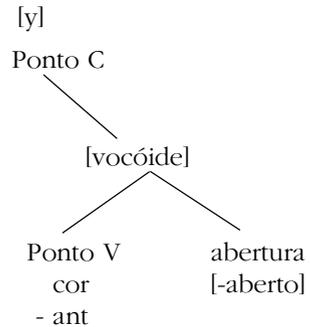
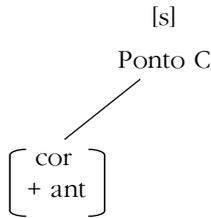
Numa perspectiva diacrônica, os fatos de não haver no latim a *fricativa palatal* e de ter ela, no português, derivado, na formação do romance, de uma [coronal] *dura* + *yode* (MATTOSO CÂMARA, 1977, p.77), permitem que se tracem as configurações desses

elementos bem como a do processo assimilatório como segue (semelhantemente a Clements, 1993):

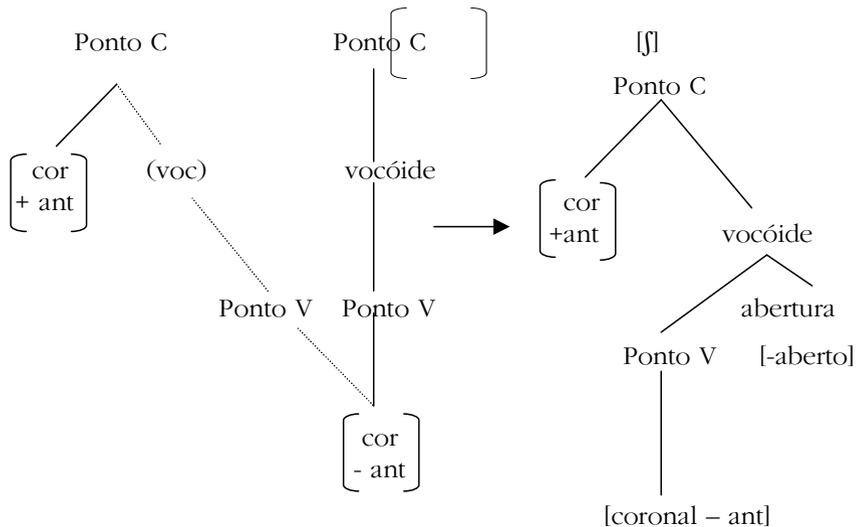
(7) assimilação da semivogal [y] pela [coronal + ant][s]

(a) Segmentos

----



(b) propagação da direita para a esquerda



Observe-se que vogal, semivogal e consoante recebem uma catalogação unificada por classe natural na proposição de Clements, datada de 1993. Ilustra-se a evolução do latim para o português, através dos exemplos:

(8) (a) *passionem* → *paixão*

(b) *vešicam* (forma hipotética) → *bexiga*

Dois outros processos variacionais muito importantes apontados por Clements (1993) referem-se ao *enfraquecimento* e *reforço* fonêmicos. O autor aborda o problema do enfraquecimento da consoante em vocóide ou a promoção dessa última categoria para consoante.

Brenner (1996), inspirada em Clements, compreende o processo variacional inserido numa escala fonêmica determinada pela abertura e pela sonoridade. Mattoso Câmara (1980, p. 61-3) sugere um sistema classificatório pautado pela abertura bucal:

(9) Categoria Grau

|           |   |
|-----------|---|
| Oclusiva  | 0 |
| Fricativa | 1 |
| Nasal     | 2 |
| Líquida   | 3 |
| Semivogal | 4 |
| Vogal     | 5 |

O quadro clássico de Mattoso Câmara assume nova relevância em estudos mais atualizados. Clements (1988, apud BRENNER, 1996, p. 86; 448) propõe uma escala da sonoridade, reformulada por Milliken, para classificar os fonemas:

(10) *Escala da sonoridade*

| O | N | L | G | V |                      |
|---|---|---|---|---|----------------------|
| - | - | - | - | + | [aberto]             |
| - | - | - | + | + | [vocóide]            |
| - | - | + | + | + | [aproximante]        |
| - | + | + | + | + | [sonante]            |
| 0 | 1 | 2 | 3 | 4 | escala da sonoridade |

O estudo baseia-se na estrutura silábica e numa teoria formal da fonologia e fonética relativa à sonoridade. A estrutura profunda de cada língua determina os princípios da organização silábica sensíveis às premissas da *sonoridade*. Distribuem-se os fonemas categorialmente, sendo, após testados quanto à posição na estrutura silábica- grau de sonoridade dos elementos situados no ataque, no núcleo e na rima. O cume silábico, nas diferentes línguas do mundo, pode ser representado por consoante ou vogal. Em português, somente a vogal se situa nessa posição. Clements, numa versão inicial, insere o traço [silábico] no limite superior da linha vertical da escala. Ocorre que Milliken (apud BRENNER, 1996) aponta para o fato de que o traço *silábico* não se mostra condizente com as propriedades físicas de um fonema, critério básico para delimitação das outras categorias. Optou, pois, pelo parâmetro da *abertura*. Verifique-se, na linha horizontal, que, nas demarcações extremas, a *obstruinte*, oclusiva e fricativa, se contrapõe à *vogal*. As categorias intermediárias são preenchidas pela *nasal*, *líquida* e *glide*. No eixo vertical, o traço [aberto] (inicialmente [silábico]) introduz uma escala decrescente de abertura e sonoridade.

O português coloquial e dialetal apresentam um processo clássico de enfraquecimento da lateral coda [ɬ], velar, em semivogal [w]. Um *corpus* muito rico representativo da fala dos pescadores e rendeiras das praias de Florianópolis foi registrado em diversas pesquisas feitas pela autora deste artigo e por alunos da graduação do curso de Letras de UFSC, coordenados pela Professora. Para ilustração, bastam os seguintes exemplos:

- (11) (a) [ˈfɛw] (fel)  
 (b) [braˈziw] (Brasil)  
 (c) [aniˈmaw] (animal)  
 (d) [ɐˈnɔw] (anzol)

Pela escala de abertura de Mattoso Câmara, a lateral portadora de grau 3, quando substituída por uma semivogal, assume a abertura 4. A escala de sonoridade de Clements qualifica a líquida como [aproximante][sonante]. Ao se realizar como um elemento com o

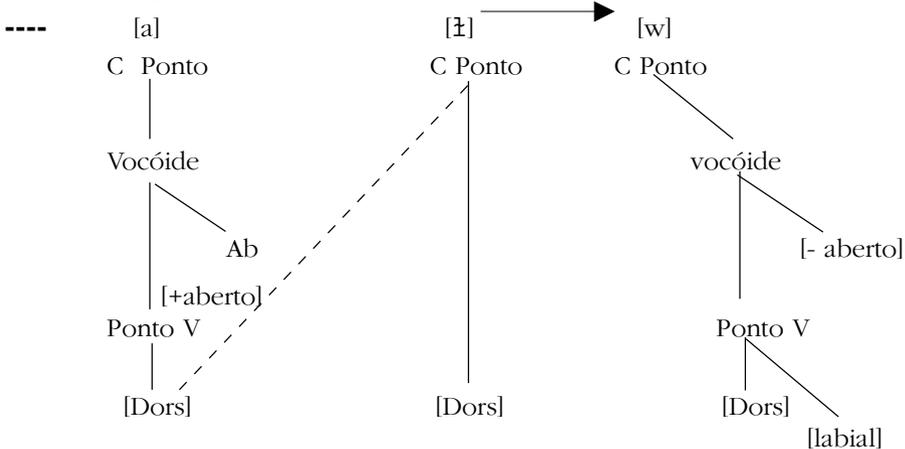
traço [vocóide], ou seja, com a propriedade vocálica, na categoria de semivogal, aumenta seu grau de sonoridade e abertura.

Toda sílaba possui um ciclo de sonoridade caracterizado por um auge e um declive (CLEMENTS, 1988, apud BRENNER, 1996). Situando-se a lateral no declive, posição de coda, ocorre um decréscimo da abertura e da sonoridade e ela se torna mais sujeita a transformações. Por outro lado, a líquida se classifica como [+aproximante, +sonante, -vocóide]. Representa, pois, o elemento mais aberto e mais sonoro do quadro consonântico. Conseqüentemente, pode se articular, como coda [velar], com a vogal antecedente, assimilando o traço vocálico. Em decorrência, atualiza-se como semivogal [w] [lábio-velar], segmento posterior. Esquematiza-se o processo:

(12) assimilação vocálica:

[aɿ] → [aw]

(13) propagação da esquerda para a direita:



Quanto à questão centrada neste artigo, verifica-se que a variação ocorreu no plano do modo de articulação entre categorias dispostas, na escala da abertura e sonoridade, hierarquicamente, em linhas diferentes.

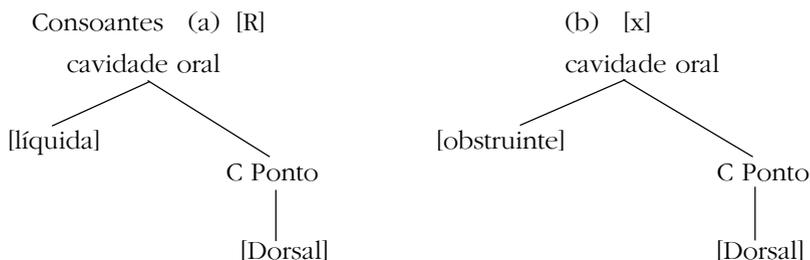
Já se aludiu anteriormente ao fato de que Brenner (1996) entendeu que o enfraquecimento e o reforço podem ocorrer num processo inter e intracategorial no sistema consonântico. Um

exemplo clássico de *reforço* fonêmico pode ser ilustrado, pois, pela variação dialetal da líquida vibrante [R]. No processo histórico, o fonema se introduziu no Brasil como [líquida] [alveolar]. Em Santa Catarina se caracteriza como [vibrante velar]. No entanto, ele pode, em variantes regionais, se atualizar como uma [fricativa], assumindo diferentes pontos de articulação, desde velar até glotal. Compare-se o falar catarinense com o carioca, inscrito o último em múltiplos manuais de fonologia. Callou & Leite (1995) fazem referência ao processo de reforço e citam dois registros fricativos para a vibrante culta do Rio de Janeiro – velar surda e glotal surda. Não mencionam a variante sonora, registrada por outros autores, resultante da assimilação do traço de sonoridade da consoante seguinte:

|                     |                           |
|---------------------|---------------------------|
| (14) Santa Catarina | Rio de Janeiro            |
| (a) [ 'maR ]        | [ 'max ]    [ 'mah ]      |
| (b) [ 'Remu ]       | [ 'xemu ]    [ 'hemu ]    |
| (c) [ 'kɔRdɛ ]      | [ 'kɔɣdɛ ]    [ 'kɔfidɛ ] |

Verifica-se a vibrante, por natureza líquida, deslocar-se para o nível da fricativa, na variação regional. Do grau 2 de sonoridade, passa para o grau 0, conforme quadro de Clements apresentado anteriormente. Assinala-se, igualmente, diminuição do grau de abertura. O reforço se efetiva no plano de modo de articulação, ocorrendo o processo entre linhas dispostas hierarquicamente, quanto à abertura e sonoridade. Observa-se também a flutuação do ponto de articulação dentro da categoria líquida ou fricativa. Ilustra-se através da configuração dos elementos implicados:

(15) realização de [R] ou [x]



Outro reforço atestado na categoria [líquida] reporta-se à variação entre elementos subcategoriais. Brenner (1996), analisando a *líquida*, interpreta a subcategoria [lateral] como mais sonora e mais aberta e a [vibrante] como a menos sonora e menos aberta, ocupando a [batida] posição intermediária.

No dialeto dos pescadores de Florianópolis, constata-se a realização da coda, [lateral, velar] como [vibrante, velar]. Sejam, entre numerosos exemplos de reforço fonêmico intracategorial (líquida):

| (16) Lateral    | Vibrante    |
|-----------------|-------------|
| (a) [a 'paɫpɛ]  | [a 'paɾpɛ]  |
| (b) [kaɫ 'sadu] | [kaɾ 'sadu] |
| (c) ['boɫsɛ]    | ['boɾsɛ]    |
| (d) [maɫ 'vadu] | [maɾ 'vadu] |

Menos freqüente, mas também acusado é o fenômeno oposto de *enfraquecimento* – a coda [vibrante] tem como segmento correlato, na variação livre, a [lateral +velar]. No plano do modo de articulação [líquida], o falante pode optar, numa variação livre, entre linhas subcategoriais hierárquicas quanto à abertura e sonoridade.

A variação dialetal implica, ainda, a transparência de certos segmentos consonânticos à assimilação vocálica. Estudos de Donca Steriade, 1986, e Carole Paradis & Jean-François Prunet (1989, apud BRENNER, 1996) mostram que a [coronal] representa a classe não-marcada quanto a ponto de articulação e que, portanto, se comporta como transparente à propagação vocálica, inclusive à propagação à distância. Os dois últimos autores prevêm que a mencionada transparência se explique pela ausência de nó categorial no elemento não-marcado.

No dialeto dos pescadores de Florianópolis, verifica-se que essa transparência se aplica muito comumente a um elemento marcado, segundo as teorias acima mencionadas: o [labial] e o [dorsal].

Observa-se a assimilação da semivogal do ditongo por vogal de sílaba precedente. Pode ocorrer cópia da semivogal com ou sem apagamento desse elemento na sílaba originária como em (17)(a). Apontam-se elementos do *corpus*:

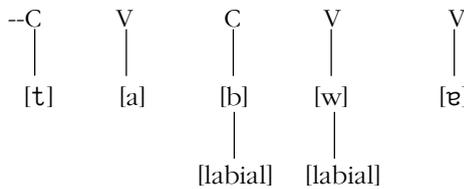
(17) Registro *Standard*

Variante regional

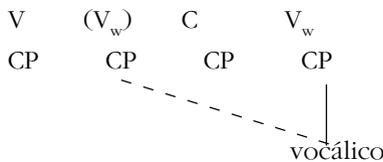
|                |                           |
|----------------|---------------------------|
| (a) [ 'tabwɛ ] | [ 'tawbɛ ] ou [ 'tawbwɛ ] |
| (b) [ko 'mew]  | [kow 'mew]                |
| (c) [ʃo 'vew]  | [ʃow 'vew]                |
| (d) [ 'magwɛ ] | [ 'mawgwɛ ]               |
| (e) [ 'nɔdwɛ ] | [ 'nowdwɛ ]               |
| (f) [ko 'low]  | [kow 'low]                |

Reforça-se o fato de que a transparência consonântica não se efetiva conforme a previsão: presença de uma categoria não-marcada. O fenômeno merece, pois, uma reflexão. Analisando a propagação transcategorial em (a) acima, percebe-se que a semivogal expande-se sem que a consoante *l*he ofereça uma real barreira. Segundo premissas de assimilação, a categoria marcada [labial] deve ser portadora de nó referente a ponto de articulação e deve, por conseguinte, se comportar como opaca à assimilação transcategorial. Configura-se abaixo a organização dos elementos de (17) (a):

(18) distribuição categorial de consoantes e vogais:



(19) propagação transcategorial



Um contexto formado pela adjacência de dois elementos [labial], um consonântico e o outro vocóide (semivogal [w]), torna plausível a explicação de que a consoante intervocálica inicial de

sílaba se mostre transparente à expansão do elemento semivocálico para a esquerda, perdendo, o segmento consonântico, o nó de ponto de articulação. A presença do referido nó implicaria cruzamento de linhas violando o princípio universal de não-cruzamento – *PCO* (CLEMENTS, 1993, p. 134-5). A cópia da semivogal na primeira sílaba originando um ditongo, determina ou não o apagamento, na segunda sílaba, do elemento copiado. Portanto, a consoante [labial], no contexto em estudo, não se comporta como neutra, opaca, à assimilação transcategorial.

Entende-se, outrossim, que o elemento vocálico da primeira sílaba deva formar um contexto favorável à assimilação e à ditongação. Tanto em (17)(a) como (d) os ditongos se estruturam em torno de [a]. Por outro lado, as sílabas de (b), (c) se forjam tendo como núcleo vogais mediais. Em (e) e (f) está prevista a transparência do nó [coronal] como categoria não-marcada.

No dialeto em questão, a dissimilação também não se restringe à transparência da *coronalidade*. As consoantes [labial] e [velar] facultam, igualmente, o processo. Sejam:

|                                |                    |
|--------------------------------|--------------------|
| (20) Português <i>Standard</i> | Português dialetal |
| (a) ['lɛɫpɔdɛ]                 | ['lɛɫpɪdɛ]         |
| (b) ['ʃakɔrɛ]                  | ['ʃakɪrɛ]          |

A análise merece maior estudo e aprofundamento que parecem, no entanto, não adequados à natureza deste trabalho.

Conclui-se, pois, que o sistema dialetal compreende regras peculiares, parâmetros próprios, condizentes, porém, com princípios universais do componente fonológico.

#### 4. CONSIDERAÇÕES FINAIS

Este trabalho demonstrou, através da análise de alguns fenômenos dialetais do português do Brasil, que o processo variacional dessa língua se insere em princípios universais da fonologia multilinear. O conceito de classe natural formulado em

Praga inspirou os estruturalistas e, na atualidade, fonólogos defensores da geometria dos traços, entre eles, George Clements. O conceito de classe natural e o de hierarquia categorial facultaram a explicação da variação livre no Brasil entre categorias do mesmo nível ou do mesmo plano, bem como entre categorias dispostas hierarquicamente. Tornou-se, pois, possível ilustrar o enfraquecimento e o reforço fonêmicos, bem como aludir a processos de assimilação transcategorial da língua em questão.

## REFERÊNCIAS BIBLIOGRÁFICAS

BRENNER, T. de M. *Une approche multilinéaire de la variation dialectale des consonnes occlusives et liquides chez les pêcheurs de Florianópolis*. 1996. Tese (Doutorado) – Sorbonne-Nouvelle, Paris.

BRENNER, T. de M. *Fundamentos da fonologia multilinear*. Projeto de Pesquisa. Florianópolis, UFSC, em andamento.

CALLOU, D.; LEITE, Y. *Iniciação à fonética e à fonologia*. Rio de Janeiro: Zahar, 1995.

CAMARA JR., M. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Padrão, 1977.

CAMARA JR., M. *Princípios de lingüística geral*. Rio de Janeiro: Padrão, 1980.

CLEMENTS, G. The geometry of phonological features. *Phonology Yearbook*, n. 2, p. 225-52, 1985.

CLEMENTS, G. Lieu d'articulation des consonnes et des voyelles: une théorie unifiée. In: LAKS, B.; RIALLAND, A. (Ed.). *Architecture des représentations phonologiques*. Paris: CNRS Ed. 1993. p. 101-145.

# Simbolismo sonoro do PB: o estudo dos ideofones

Regina Célia Fernandes Cruz  
Universidade Federal do Pará

Helane de Fátima Gomes Fernandes  
Universidade Federal do Pará

## Abstract

[Introduction] The studies about ideophones in Brazilian Portuguese (PB) reveal that in the Portuguese system the ideophones are typically formed by bilabial sounds (CRUZ, 2000a). [aim] The main aim of this study is to verify if this is a typical feature of the PB or a resource of the sound symbolic expressivity. [methodology] In this study we observed the ideophones phenomenon in three kinds of varieties of Portuguese: (i) Afro-brazilian – ABP (CRUZ, 2000a), Brazilian dialect spoken in Amazonian by the maroon community; (ii) Amazonian Portuguese – AM (TRINDADE, 1992; CRUZ, 2000a; CASSIQUE, 1995) and (iii) standard Brazilian Portuguese – PB (OLIVEIRA, 2000; BORGES, 1991). We considered other data too: a) data collected through direct observation of different conversational situations and b) the data selected from cartoons. Particularly, we will analyse two aspects: the phonetic feature of the sounds found in the ideophones and the association of the meanings established. [results] The study of the Brazilian Portuguese ideophones show us: a) a tendency to bilabial sounds, b) a tendency to expression links to emission of the sound and c) ideophones are not dialectal features of ABP. [conclusion] The results indicate a necessity to study more the PB.

## 1. INTRODUÇÃO

O fato de o simbolismo sonoro e os ideofones serem processos que estão presentes na complexa comunicação humana é fator suficiente para que tais fenômenos mereçam ser analisados com maior profundidade.

Apesar de ser um assunto marginal na literatura, os trabalhos realizados sobre o fenômeno em outras línguas conseguem avançar bastante na definição das características gerais e típicas deste fenômeno. Entretanto, pouco se tem trabalhado sobre o referido processo no português brasileiro (doravante PB). Face a essa carência, o presente trabalho se constitui em um aprofundamento da descrição do processo de formação de ideofone no PB presente em Cruz (2000a).

Das diversas definições encontradas na literatura sobre ideofones, neste trabalho destacam-se as seguintes:

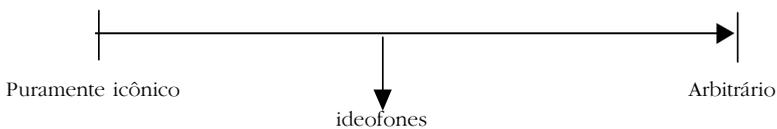
[...] forms which can function as complete sentences and whose morphemic constituents are phonic features. (DIFFLOTH, 1972)

I recognize two characteristic properties: phonetic play reflecting sound symbolism, and reduplication as a structural property. (CRUZ, 2000a)

Phonetic play is the utmost intrinsic property of ideophones by which iconic relationship between the referent and the meaning can be achieved. Reduplication is the structural property that ideophones often exhibit. (LEE, 1992)

Apesar de Lee (1992) colocar em evidência o jogo fonético e a reduplicação como características importantes na formação de ideofones, a reduplicação não é considerada no presente trabalho como característica essencial na sua estrutura por dois fatores: (i) o fato de a reduplicação se fazer presente na formação das onomatopéias e (ii) o fato de o *corpus* utilizado apresentar dados cujo jogo fonético foi obtido sem a reduplicação.

Conseqüentemente, neste trabalho enfocar-se-ão muito mais as características fonéticas dos ideofones. Tal aspecto fonético é reforçado na escala de Fordyce (1998, apud LEE, 1992). Fordyce (1988) estabeleceu uma escala na qual as palavras motivadas de um forte simbolismo sonoro formam um *continuum* entre as categorias que são tipicamente icônicas e aquelas que não o são, explicitados pelo esquema abaixo.



Bartens (2000) apresenta outras características atribuídas aos ideofones afirmando que são sons: a) onomatopéicos acompanhados de gestos; b) presentes na comunicação oral em situação de fala informal; c) com tendência a desaparecer quando a padronização de uma língua ganha espaço e; d) típicos da comunicação humana, funcionando como uma sentença completa ou um termo dentro da mesma, ocupando a posição de um substantivo, de um verbo, advérbio, adjetivo e outros componentes que marcam a situação conversacional (DOKE, 1935). As características apontadas por Bartens (2000) para definir os ideofones expressam ainda mais a íntima relação entre ideofones e o simbolismo sonoro de uma língua (LEE, 1992 e BARTENS, 2000). Todas as características apresentadas por Bartens (2000) foram levadas em consideração ao se proceder à análise dos dados.

Não se pretende aqui abranger todos os aspectos possíveis de seu enfoque, mas fazer uma descrição das características dos ideofones no PB. Espera-se, também, contribuir para estudos posteriores, visto que há ainda muito que se descobrir sobre a relação ideofones e simbolismo sonoro.

## 2. PROBLEMATIZAÇÃO

Com o objetivo de identificar herança do substrato africano no português brasileiro, Cruz (2000a), em seu estudo sobre o português

falado por comunidades quilombolas do Pará (doravante ABP), aponta como marca de identidade lingüística das referidas comunidades os ideofones. Os ideofones estudados por Cruz (2000a) como típico do ABP são os presentes na tabela 1

Tabela 1- Exemplos de ideofones do ABP (cf. CRUZ, 2000a)

| Ideofone | Exemplo                                                           |
|----------|-------------------------------------------------------------------|
| [tətɛl]  | – eu pensava que ia morrer, falava coisa teteé.                   |
| [teitei] | – Ele está teitei de açai.                                        |
| [te'ite] | – Vê se eles gostem que a gente tire o deles. Ela ainda ri teité. |
| [nn]     | – começa a se espalhar né? - hum, hum.                            |
| [hh]     | – em compromisso? - hem, hem”.                                    |

Baseando-se em estudos que apresentam os ideofones como típicos das línguas africanas – (BARTENS, 1997; CHILD, 1994; WESCOTT, 1977; SAMARIN, 1965; NEWMAN, 1968) – os ideofones foram vistos pela autora, em um primeiro momento, como herança direta das línguas africanas pelo seu nítido processo de reduplicação (CRUZ, 2000b).

Entretanto, os ideofones não são apenas típicos das línguas africanas, eles fazem parte de línguas como coreano (CRUZ, 2000b; LEE, 1992), línguas indo-pacíficas munda (KUIPER, 1965), línguas indianas (EMENEAU, 1969), japonês (ALFONSO, 1966 e TAKEMOTO, 1998), algumas línguas indo-arianas (DIMOCK, 1957) e chinês. Constituem também a prova de influência direta do chinês no inglês falado em Singapura (LIM, 1998). Os mesmos ideofones identificados como típicos do ABP por Cruz (2000a) apresentam similaridades acentuadas de forma, função e significação com ocorrências encontradas em línguas indígenas da família tupi.

Assim como se atesta que os mesmos ideofones estudados por Cruz (2000a) pertencem ao sistema lexical e gramatical de línguas indígenas que sofreram forte influência da língua geral como tembé (BORGES, 1991), ou da própria língua geral amazônica (ARAÚJO, 1989), atesta-se também que eles pertencem a outras línguas que não fazem parte da família tupi-guarani como o guató (PALÁCIO, 1984), em que é possível encontrar a partícula afirmativa do ABP [NN]. Da mesma

forma o qualificador [tete'ɛ] apresenta ocorrências muito próximas no Tembé (EIRÓ, 2001). Ver tabela 2.

Tabela 2 – Exemplos de ideofones em línguas indígenas

| Ocorrências                | Transcrição                                        | Função               | Línguas                              | Fontes                                                        |
|----------------------------|----------------------------------------------------|----------------------|--------------------------------------|---------------------------------------------------------------|
| Hemhem                     | [hh]                                               | Partícula afirmativa | Kamauirá                             | (SEKI, 2000)                                                  |
| Humhum                     | [ɥɥ]                                               | Partícula afirmativa | Guató                                | (PALÁCIO, 1984)                                               |
| Teéteé,<br>Teteé,<br>Teuté | [teʔé:teʔé],<br>[tete'ɛ],<br>[teu'tɛ],<br>[te'tɛʔ] | Intensificador       | LGA, Ka'apor,<br>Makurapi<br>e Tembé | (CALDAS, 2001;<br>BRAGA, 1992;<br>BORGES, 1991<br>EIRÓ, 2001) |

Do trabalho de Cruz (2000a), retiveram-se dois aspectos para revisão e aprofundamento:

1. A preferência do PB pelo traço bilabial na formação de seus ideofones;
2. os ideofones enquanto marca dialetal do ABP.

Resolvemos, portanto, aprofundar e avançar um pouco mais na descrição feita por Cruz (idem), buscando, desta vez, verificar as seguintes hipóteses:

- a) O português brasileiro tem preferência pelo traço bilabial?
- b) Em qual função gramatical há maior ocorrência de ideofones?
- c) Qual tipo de significado é preferencialmente expresso pelos ideofones?
- d) Os ideofones funcionam como marca dialetal?
- e) Os ideofones funcionam como identidade lingüística do ABP?
- f) Há alguma relação direta entre ideofones e língua tonal?
- g) Os ideofones típicos do ABP seriam herança da língua geral amazônica?
- h) Os ideofones seriam típicos de situações de fala informal, como conversas informais e narrativas orais?

### 3. METODOLOGIA

Para comprovação das hipóteses acima listadas, procedeu-se da seguinte forma.

Primeiramente, recorreu-se a um levantamento dos dados nos *corpora* consultados, ou seja: em Bartens (2000), em Cruz (2000a), em Cassique (1995) e em Oliveira (2000).

Ampliou-se o *corpus* buscando verificar a ocorrência de ideofones na fala cotidiana, em particular em conversa informal, por anotação direta e sem utilização de gravador cassete. Coletaram-se, também, dados de revistas em quadrinhos infantis, com o objetivo de verificar se a produção de ideofones é típica de uma situação de fala em particular.

O universo pesquisado foi completado consultando-se trabalhos referentes a línguas indígenas brasileiras, dando preferência a estudos de descrição de línguas pertencentes à família tupi, uma vez que tudo indicava serem os ideofones típicos do ABP uma marca de influência da língua geral amazônica. Neste sentido, buscou-se um levantamento de dados em Araújo (1989), Braga (1992), Caldas (2001), Eiró (2001), Julião (1983), Maciel (1991), Rodrigues (1995) e Silva (1999).

Após o levantamento dos dados, passou-se para a etapa de sistematização dos dados e, para esta fase, utilizou-se o programa Excel. Uma planilha Excel possibilitou uma melhor visualização e uma melhor organização dos dados. Na planilha, os dados foram organizados, constando das seguintes informações: transcrição fonética, função gramatical, significação, variedade lingüística, situação de fala, fonte consultada, contexto e traço fonético dominante. Encontra-se, em anexo, um quadro (quadro 1) com os dados dos ideofones originais de Cruz (2000a, b). Uma vez tendo os dados sido organizados na planilha Excel, foi possível passar à análise quantitativa e qualitativa. Foi feita, primeiramente, uma triagem que tinha como objetivo observar a frequência de ocorrências das características acima mencionadas por dado. Todos os resultados obtidos na análise foram organizados em gráficos para melhor visualização dos dados.

#### 4. CORPUS FINAL

Ao todo, identificaram-se 85 ideofones cuja distribuição por variedade lingüística do português brasileiro é a seguinte: (i) ABP (19 dados), (ii) AM-português da Amazônia (21 dados) e (iii) PB-português brasileiro padrão (45 dados).

Obtiveram-se também como material de comparação dados das principais línguas indígenas: Amanajá, Anambé, Kamaiurá, Ka'apor, Makurap, Makú, Parakanã, Tembé, Guató. Dessas línguas, obtiveram-se 29 ocorrências que chamaram atenção por sua forte semelhança de estruturação com os ideofones estudados, o que possibilitou continuar acreditando que os ideofones possuem uma forte influência das línguas indígenas no PB.

Portanto, 114 ocorrências de ideofones foram analisadas neste trabalho.

#### 5. ANÁLISE

A organização dos dados nas planilhas permitiu um melhor tratamento dos dados coletados. O fator de maior controle efetivado foi a frequência de ocorrência dos ideofones no *corpus* a partir de cada aspecto determinado como caracterizador dos ideofones (v. final do item 2). Obtiveram-se os seguintes resultados.

##### 5.1. Função gramatical

Como já mencionado anteriormente, os dados coletados foram analisados sob vários aspectos; um desses aspectos está ligado à função gramatical dos ideofones. Esse aspecto foi cogitado em decorrência de uma das definições dos ideofones afirmar que estes são combinações sonoras que podem assumir funções na frase, recaindo sobre si, por exemplo, o papel de um verbo, de um designativo ou de um qualificador e assim por diante (BARTENS, 2000; DOKE, 1935). A seguir, alguns exemplos das diferentes funções morfossintáticas que podem ser assumidas por um ideofone.

**[tʃũ]**

Apresenta função equivalente a de um verbo por expressar o ato de pular ou jogar-se.

(126) “... esses dois parceiros se jogaram na água aí ele **tchum** atrás.”

(127) “...pulou n’água ele, doido **tchum** atrás.”

**[teitei]**

Aparece freqüentemente nos dados com função de intensificador, tem o significado de ‘cheio’, ‘muito’ ou ‘bastante’.

(138) “– Ele está **teitei** de açáí .”

**[ŋŋ]<sup>1</sup>**

O ideofone ‘*bumbum*’ trata-se de uma partícula afirmativa.

(51) Locutor 1: “– é uma desculpa pra preguiça dele né tia?”

Locutor 2: “– **bumbum**.”

**[hh]<sup>2</sup>**

O ideofone ‘*bembem*’ constitui uma outra partícula afirmativa.

(45) Locutor 1: “– foi por causa do marido?”

Locutor 2: “– **bembem**.”

**[tei'te]**

O ideofone ‘*teité*’ é um qualificador, cujo significado é ‘coitado’ ou ‘infeliz’.

(118) “– Vê se eles gostem que a gente tire o deles. Ela ainda ri **teité**.”

Em decorrência desses papéis assumidos pelos ideofones, resolveu-se pesquisar que função gramatical é melhor e é mais representada na frase pelos ideofones do PB. O gráfico 1 ilustra os resultados alcançados.

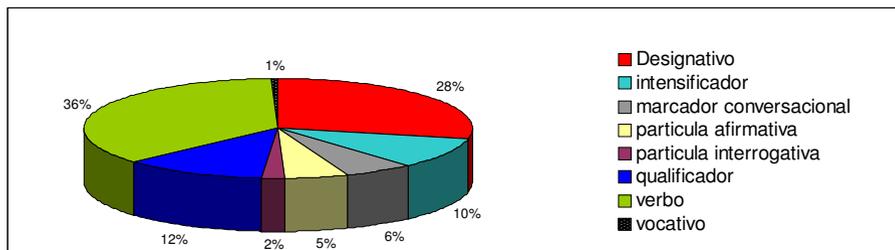


Gráfico 1 – Distribuição dos ideofones em funções gramaticais (total de 85 dados)

O gráfico 1 mostra que as funções gramaticais mais expressivas foram as de designação, intensificação, marcador conversacional, partículas afirmativas, partículas interrogativas, qualificadores e de verbos. Porém, de todas essas funções, o papel de verbo é o melhor representado pelos ideofones tanto no aspecto qualitativo como no quantitativo, chegando ao percentual de 36% das ocorrências. Tal resultado é ratificado pelo gráfico 1 que mostra os ideofones como melhores representantes do campo semântico de ações, e, principalmente, ações ligadas à emissão de sons, como é possível ver mais adiante. Em segundo lugar, aparece a função de designativo com um bom percentual (28%). As outras funções que aparecem com percentual significativo e que também estavam sendo esperadas são as funções de qualificadores (12%) e intensificadores (10%).

O próximo aspecto a ser observado diz respeito ao aspecto sociolinguístico. Procurou-se verificar qual variedade do português faz mais uso de ideofones como recurso linguístico.

## 5.2. Variedade linguística

A observação da distribuição dos ideofones nas variedades linguísticas do português e em outras línguas brasileiras é importante para detectar se o uso de ideofones se dá em determinada variedade do português ou se é um traço linguístico universal. O gráfico 2 reflete os resultados obtidos.

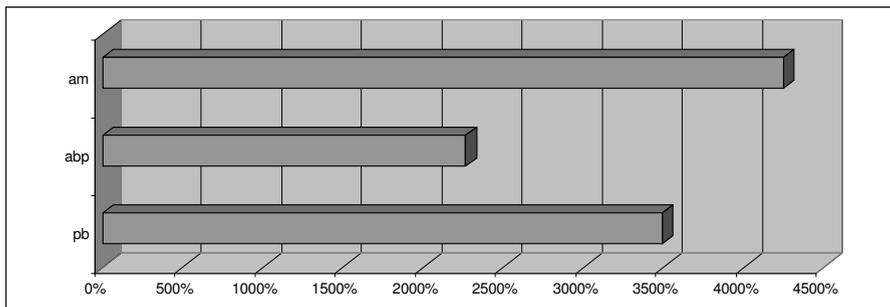


Gráfico 2 – Distribuição dos ideofones por variedade lingüística do português (total de 85 dados)

Os resultados da distribuição de ideofones nas variedades lingüísticas são claros em informar que, segundo a freqüência de ocorrências, o português regional da Amazônia (AM), com cerca de 42% das ocorrências, apresenta um maior uso de ideofones na sua comunicação. Observa-se também que no português brasileiro padrão (PB) se faz bastante uso de ideofones como recurso de expressividade simbólica 34%. Esta primeira distribuição não apresenta os ideofones como uma marca dialetal de nenhuma variedade do português em particular. Abaixo, eis alguns exemplos de ideofones do português por variedade lingüística.

| Variedade Lingüística | Contexto                                | Fonte             |
|-----------------------|-----------------------------------------|-------------------|
| <b>Am</b>             | – agora <i>malamá</i> a vespera.        | (CAMPOS, 1998)    |
| <b>ABP</b>            | – Mapará <i>fifiti</i> .                | (CRUZ, 2000a)     |
| <b>PB</b>             | – Eles estão tendo um <i>teretetê</i> . | Observação direta |

### 5.3. Traço bilabial geral

Cruz (2000a) detectou que, no PB, existe uma preferência pelo uso de traço bilabial na formação de seus ideofones. Os exemplos a seguir ilustram a preferência por esse traço.

Contexto

Transcrição

(09) –...ele só deu um pulo, **bam** e **bacu** no defunto.

[bã], [baku]

(13) –...eu não sei o quê e **bebebê**, e **bababá**.

[bebe'be], [baba'ba]

(76) – taquei-lhe na cara dele, **pá** e eu **bacu**.

[pa], [baku]

(33) – a minha teve nove filhos, um morreu **gito** logo que nasceu. [ʒitʊ]

(116) – Ele virou agora um **teba** de homem.

[teβ]

O gráfico 3 mostra essa disposição no português

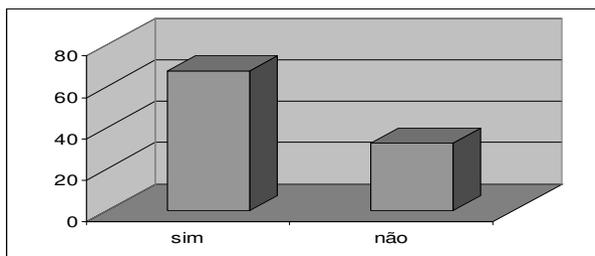


Gráfico 3 – Frequência da ocorrência da articulação bilabial no conjunto dos dados (total de 114 dados)

Os dados obtidos e analisados confirmam a hipótese de Cruz (2000a) de que existe uma preferência maior pelo traço bilabial na formação da expressividade sonora do PB.

Para conhecer melhor essa preferência pelo traço bilabial na formação dos ideofones no português brasileiro, buscou-se saber se esta preferência permanece em todas as variedades lingüísticas ou se existe alguma que foge a essa regra. Observe o gráfico 4.

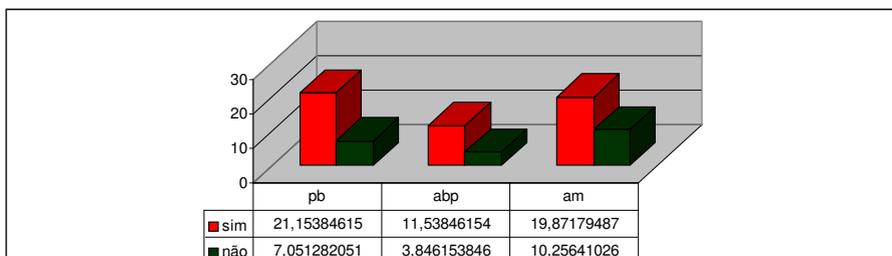


Gráfico 4 – Ocorrência de articulação bilabial nos ideofones por variação lingüística (total de 156 dados)

Os dados confirmam a presença do traço bilabial como preferência no português brasileiro, principalmente pelo PB, que apresentou um percentual bastante significativo de uso (21%). Em segundo lugar tem-se o AM (19%).

No caso do ABP, Cruz (2000a) já havia detectado que não existe essa preferência por traço bilabial nessa variedade do português em particular.

Percebe-se uma aproximação visível entre as línguas indígenas e o português brasileiro nos resultados a respeito da utilização de traço bilabial na formação dos ideofones. Ver o gráfico 5.

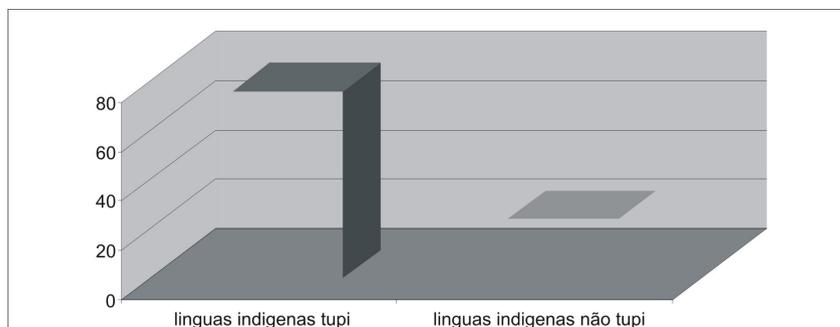


Gráfico 5 – Freqüência de articulação bilabial nas línguas indígenas

Os resultados mostram que as línguas indígenas de origem tupi também têm essa preferência pelo traço bilabial na formação de seus ideofones.

#### 5.4. Situação de fala

A ocorrência em diferentes situações de fala foi outro aspecto importante a ser observado, pois assim fica claro se os ideofones do PB, assim como os de outras línguas, apresentam preferência por uma determinada situação de fala. Fazendo a distribuição dos dados, alcançaram-se os seguintes resultados como mostra o gráfico 6.

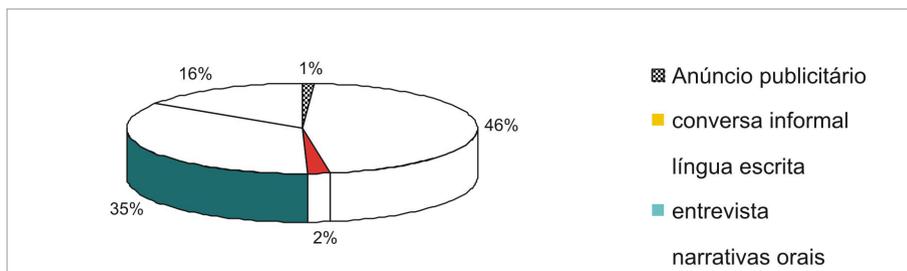


Gráfico 6 – Distribuição dos ideofones em situação de fala

Os estudos anteriores delinearão características dos ideofones em várias línguas como o coreano (LEE, 1992), japonês (TAKEMOTO, 1998), crioulos do atlântico (BARTENS, 1997), africano (CHILD, 1994) e inglês (WESCOTT, 1977). Entre as já citadas está evidente a preferência de utilização de ideofones na comunicação oral. No gráfico 6 percebe-se que os ideofones do PB apresentam essa mesma característica, pois o fenômeno foi pouco encontrado em língua escrita que tem apenas 2% dos dados nesta modalidade. Tal resultado explica o motivo pelo qual a observação dos dados em revistas em quadrinhos não aparece no trabalho, pois não foram detectadas ocorrências suficientes para que participassem do *corpus* utilizado. A maior ocorrência de ideofones se deu na situação de conversa informal com 46% das ocorrências. As entrevistas e as narrativas orais aparecem com 35% e 16% das ocorrências, respectivamente.

### 5.5. Campo semântico

Ao proceder-se a um levantamento do campo semântico melhor representado pelos ideofones, alcançou-se um nível bastante significativo de compreensão dos mesmos, e até uma melhor descrição (ver o gráfico 7). A seguir são mostrados alguns dos campos semânticos abrangidos pelos ideofones.

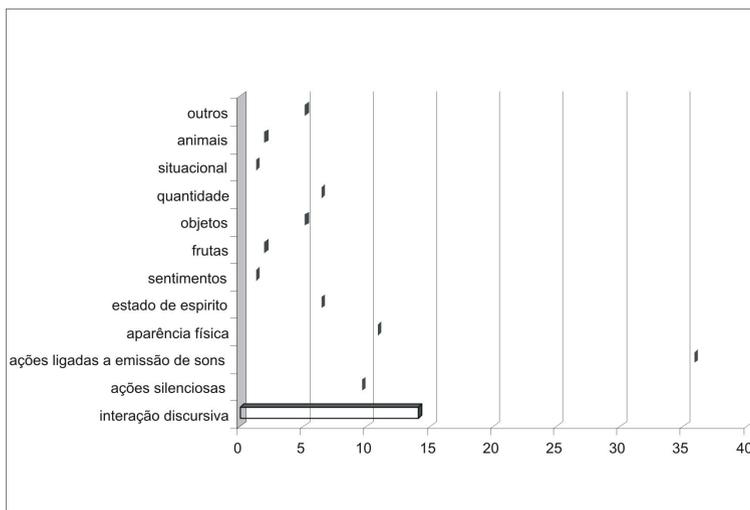


Gráfico 7 – Preferência do campo semântico dos ideofones

### [dridridri]

Esse é um ideofone que está relacionado semanticamente com ações ligadas a emissão de sons

(31) “...ele com o defunto na...no ombro e *dridridri* e ele atrás.”

### [pa]

O ideofone *pa*, nesta frase, está ligado semanticamente a interações discursivas.

(77) “...*pá* conversa vai *pá* conversa vem.”

### [zito]

O ideofone ‘*jito*’ está relacionado à idéia de intensificação, de pequeno.

(40) “Eu vi um passarinho *jito, jito*.”

### [teb]

O ideofone ‘*teba*’ também está relacionado à idéia de intensificação, mas como aumentativo.

(129) “Ele virou agora um *teba* de homem.”

O gráfico 7 demonstra que, preferencialmente, os ideofones optam por representações de ações; tal fato anda em harmonia com os resultados obtidos no gráfico 7. Porém, uma observação importante é que essas ações preferenciais estão ligadas à emissão de um som. Esta observação nos faz relacionar o som emitido por uma ação à estrutura e ao significado expresso pelo ideofone.

Nos dados coletados para o presente trabalho, é possível perceber também que existe uma semelhança entre os sons utilizados na estruturação desses dados e os sons tradicionalmente ligados à idéia de bater, cair, pular, etc. (ver tabela 3). E tal fato mostra que os sons produzidos por essas ações estão representados na estruturação dos ideofones majoritariamente estudados. A relação direta entre a idéia do som produzido pela ação e os sons utilizados para sua representação nos ideofones expressa uma relação icônica e, logo, simbólica. Sendo o simbolismo intimamente ligado à significação, eis a razão de uma relação entre ideofones e simbolismo sonoro.

Tabela 3 – Ideofones representando ações ligadas à emissão de som

| Ideofone | Exemplo                                           | Significação |
|----------|---------------------------------------------------|--------------|
| [pal]    | – taquei- lhe na cara dele, <b>pá</b> e eu bacu.  | Bater        |
| [ˈbaku]  | – taquei- lhe na cara dele, pá e eu <b>bacu</b> . | Bater        |
| [tʃãw]   | – arrastou na cabeça da velha pá, <b>tchão</b> .  | Cair         |
| [tʃũ]    | – ...pulou n'agua ele, <b>tchum</b> atrás         | Pular        |

Outro resultado muito importante, e já esperado, diz respeito à interação discursiva preferida pela maioria dos ideofones também confirmado pelo gráfico 6, que mostra uma maior presença de ideofones na conversa informal.

## 6. RESULTADOS

Tomando como base a análise feita no item anterior, os dados mostram as seguintes respostas às perguntas do item 2:

- a) o português brasileiro tem preferência pelo traço bilabial na formação de seus ideofones. Os resultados confirmam a hipótese de Cruz (2000a) que já indicava como tendência do PB a utilização do traço bilabial na formação dos ideofones;
- b) o verbo é a função gramatical mais representada por ideofones na oração;
- c) conseqüentemente, há uma preferência de ideofones na representação do campo semântico de ações relacionadas à emissão de sons;
- d) constatou-se que as ocorrências de ideofones se deram com um significativo percentual em quase todas as variedades, o que nos leva a concluir que não se trata de uma identidade lingüística do ABP, mas sim um processo que ocorre em diversas variedades do PB. A única característica peculiar dos ideofones do ABP continua sendo que os mesmos não têm preferência pelo traço bilabial;

No entanto, a presença ativa de ideofones em línguas tonais nos dá elementos para continuar investigando se a formação de ideofones em uma língua é uma característica universal ou típica de línguas tonais;

- e) como foi encontrado um bom percentual de ideofones nas variedades lingüísticas estudadas, conclui-se que os ideofones não servem como marca dialetal de uma determinada variedade do PB;
- f) os ideofones do português brasileiro encontram-se essencialmente na modalidade oral e se fazem mais presentes na fala espontânea.

## 7. CONCLUSÃO

O estudo dos ideofones do português brasileiro mostram que há: a) a preferência pelo traço bilabial na formação dos mesmos e b) a preferência pela representação de ações relacionadas à emissão de sons.

O fenômeno ocorreu proporcionalmente nas três diferentes variedades lingüísticas, o que leva a concluir que: a) os ideofones não constituem uma marca dialetal do ABP, pois detectou-se que a

característica peculiar do mesmo é a não preferência pelo traço bilabial na formação de seus ideofones; b) igualmente aos demais ideofones estudados em outras línguas, os ideofones do português brasileiro se fazem presentes na comunicação oral e, principalmente, na fala espontânea como conversa informal, entrevistas e narrativas orais.

Como assinalado na introdução deste artigo, foi possível abordar apenas alguns processos estudados no PB. A certeza deixada aqui é a da necessidade de se conhecer melhor o funcionamento do PB no que diz respeito ao seu sistema simbólico e seus recursos lingüísticos na comunicação oral.

## NOTAS

<sup>1</sup> Cruz (1994) apresenta uma descrição detalhada dessa partícula afirmativa na variedade lingüística do ABP.

<sup>2</sup> Idem.

<sup>3</sup> Para a transcrição dos dados foi utilizado o alfabeto SAMPA.

## REFERÊNCIAS BIBLIOGRÁFICAS

ARAÚJO, L. S. de. *Aspectos da língua Gavião- Jê: vocabulário*. 1989. Tese (Doutorado) - Rio de Janeiro.

ALFONSO, A. *Japanese language patterns*. Tokyo: Sophia University Press, 1966. V. 1, p. 450- 454.

BARBOSA, J. F. *Contribuição ao conhecimento da língua Amanajé*. Belém: UFPA, 2002.

BARTENS, A. *Ideofones and sound symbolism in Atlantic creoles, Abstract for the SPCL*, conference at the University of Westminster, London, June 1997, p. 26-28.

BARTENS, A. *Ideophones and sound symbolism in Atlantic creoles*. Helsinki: The Finnish Academy of Sciences, 2000. 168p.

BORGES, L. C. *A língua geral amazônica: aspectos de sua fonêmica*. 1991. 98 f. Dissertação (Mestrado) – IEL /Unicamp.

BRAGA, A. de O. *A fonologia segmental e aspectos morfológicos da língua Makurap (TUPI)*. 1992. Dissertação (Mestrado) – Unicamp, Campinas.

- CAMPOS, B.; OLIVEIRA, J. *Do Pucuru à Panelinba: dados do vocabulário mocajubense*. Cametá: UFPA, 1998.
- CALDAS, R. B. C. *Aspectos, modos de ação e modalidade na língua Ka'apor*. 2001. Dissertação (Mestrado) – UFPA.
- CASSIQUE, O. “*Os Tambouros do Itapocú*”: relatório científico do projeto Marcas Lingüísticas da fala das comunidades negras de Cametá, Belém, Pará, 1995.
- CASTILHO, A. O português do Brasil. In: ILARI, R. (Ed.). *Lingüística Românica*, São Paulo: Ática, 1992. p. 237-269.
- CHILD, G. *African Ideophones*. In: NICHOLS, L. J.; OHALA, J. (Ed.). *Studies in Sound Symbolism*. Cambridge: University Press, 1994. p.178-204.
- CRUZ, R. Sound Symbolism in Brazilian Portuguese: a study of ideophones. In: *The proceedings of the conference gesture: meaning and use*, Porto (Portugal): Universidade de Fernando Pessoa, 2000a.
- CRUZ, R. *Analyse phonologique et acoustique du portugais parlé par des communautés noires de l'Amazonie (Brésil)*. 2000b. Tese (Doutorado) – Université de Provence.
- DIFFLOTH, G. Notes on Expressive Language. In: *Proceedings of the Chicago Linguistic Society* 8, 1972.
- DOKE, C. M. *Bantu Linguistics Terminology*. London: Longmans, Green, 1935.
- DIMOCK, E. C. *Symbolic forms in Bengali*. Bulletin of the Deccan College Research Institute, 1957, v. 18, p. 22-28.
- EIRÓ, J. G. *Contribuição à análise fonológica da língua Tembé*. Belém: UFPA, 2001.
- EMENEAU, M. B. Onomatopoeics in the Indian linguistics area. *Language*, v. 45, p. 274-299, 1969.
- FORDYCE, J. F. *Studies in Sound Symbolism with Special Referent to English*. PhD dissertacion, University of California at Los Angeles, 1988.
- JULIÃO, M. R. *A língua dos índios do rio Cairari*. Belém: UFPA, 1993.
- KUIPER, F. B. J. *Consonantes variation in Munda, Indo- Pacific Linguistic Studies*, part 1, North- Holland, Amsterdam, 1965, p. 54-89.
- LEE, J-S. *Phonology and sound symbolism of Korean ideophones*. 1992. 379 f. (Dissertacions) - UMI.
- LIM, L. We talklike that Meh? No, Lah!: intonation patterns on pragmatic particles in spontaneous Singapore English- Preliminary observations. National University of Singapore, in *Sound partterns of Spontaneous Speech: productiuon and perception*. La Baume-Les-Aix, septembre, 1998.

- LEÃO, J.; GOMES, L. Um estudo semântico de palavras não dicionarizadas no linguajar cametaense: à busca de brasileirismo. Cametá: UFPA, 1998.
- MARTIN, S. E. *Phonetic Symbolism in Korean*. Uralic and Altaic Series, v. 13, Indian, 1962, p177- 189.
- MACIEL, I. *Alguns aspectos fonológicos e morfológicos da língua Máku*. 1991. Dissertação (Mestrado) – Brasília.
- NEWMAN, P. Ideophones from a syntactic point of view. *Journal of west African Languages*, Cambridge University Press, London, v. 2, p. 107-117, 1968.
- OLIVEIRA, M. *Prosodic Features in Spontaneous Narrativas*. 2000. 241f. Dissertation ( PhD.) - Simon Frase University.
- PALÁCIO, A. P. *Guatô- a língua dos índios canoeiros do rio Paraguai*. 1984. Tese (Doutorado) – UNICAMP, Campinas.
- RODRIGUES, C. L. R. *Estude morphosyntaxique de la langue Xipaya*. 1995. Tese (Doutorado) – Paris.
- SAMARIN, W. J. Perspectives on African Ideophones, *African Studies*, xxiv, p. 117-121, 1965.
- SEKI, L. *Gramática do Kamaiurá: Língua Tupi-guarani do Alto Xingu*. Campinas, SP: Editora UNICAMP; São Paulo, SP: Imprensa Oficial, 2000.
- SILVA, A. C. S. e. *Aspectos da referência alternada em Parakanã*. 1999. Dissertação (Mestrado) – UFPA, Belém.
- TAKEMOTO, E. Vocalisation des gestes: les idéophones japonais. In: SANT, S. et al. (Ed.). *Oralité et Gestualité: communication multimodale, interaction*, 1998, p. 539-542.
- TRINDADE, R. *O som da fala dos pescadores de Cametá*. 1992. Dissertação (Mestrado) - Departamento de Linguística - Universidade Federal de Santa Catarina.
- WESCOTT, R. *Ideophones in Bini and English*. Forum Linguisticum (2), 1977.

## ANEXO

## Quadro de Dados

Quadro 1 - Ideofones identificados por Cruz

| Ideofone | Transcrição Fonética | Classe gramatical            | Significação                   | Variação lingüística | Situação de fala  | Fonte                     | Exemplo de ocorrência                                                       |
|----------|----------------------|------------------------------|--------------------------------|----------------------|-------------------|---------------------------|-----------------------------------------------------------------------------|
| Fifiti   | [fɨ'fɨtɨ]            | Intensificador               | Minúsculo                      | ABP                  | Conversa Informal | Cruz (2000b)              | Mapará fifiti                                                               |
| Gito     | [ʒitu]               | Intensificador               | Diminutivo                     | ABP e AM             | Conversa Informal | Cruz (2000a, 2000b)       | Eu vi um passarinho jito, jito                                              |
| Hem, hem | [hɨ]                 | Partícula afirmativa         | Sim                            | ABP                  | Entrevista        | Cruz (2000a, 2000b, 1994) | – em compromisso?<br>– hem, hem                                             |
| Hum, hum | [ɨɨ]                 | Partícula afirmativa         | Sim                            | ABP                  | Conversa Informal | Cruz (2000a, 2000b, 1994) | – começa a se espalhar né? – hum,hum                                        |
| Malamá   | [mala'ma]            | Intensificador, qualificador | não muito, mais ou menos quase | ABP/AM/PB            | Entrevista        | Cruz (2000a, 2000b, 1996) | – eu vou pedir pra tua tia fazer um ché malamá doce.                        |
| Teba     | [teɓ]                | Intensificador,              | Aumentativo                    | ABP/ AM              | Conversa Informal | Cruz (2000b, 2000a, 1996) | – Ele virou agora um teba de homem                                          |
| Teité    | [teɨ'teɨ]            | Qualificador                 | Coitado, infeliz               | ABP                  |                   | Cruz (2000a, 2000b, 1996) | – era para apruntare essa capelinha aí, mas pobre teité num faz o que quer. |
| Teitei   | [teɨ'teɨ]            | Intensificador               | Muito, bastante, cheio         | ABP                  | Conversa Informal | Cruz (2000a, 2000b, 1996) | – Ele está teitei de açaí                                                   |
| Teteé    | [teɨ'teɨ'e]          | Intensificador               | à toa                          | ABP                  | Entrevista        | Cruz (2000a, 2000b, 1996) | – eu pensava que ia morrer, falava coisa teteé                              |

# The production of English initial /s/ clusters by Portuguese and Spanish EFL speakers

Andreia Schurt Rauber  
Universidade Federal de Santa Catarina

Barbara Oughton Baptista  
Universidade Federal de Santa Catarina

## Resumo

Este estudo revisa, verifica e explica os diferentes resultados obtidos por Carlisle (1991, 1992, 1997) e Rebello (1997) sobre a produção de encontros consonantais iniciais /sC(C)/ em inglês por falantes de português e espanhol, complementando os resultados obtidos em pesquisas anteriores na área de interfonologia, já que o mesmo *corpus* foi usado para comparar a produção de falantes das duas línguas nativas. Os resultados revelam que a adição de uma vogal (epêntese) aos encontros consonantais iniciais é a estratégia mais utilizada para se lidar com as dificuldades relativas às estruturas silábicas mais difíceis, o que pode ser resultado da relação entre a interferência da língua materna e os universais lingüísticos. Os resultados fornecem informações relevantes ao ensino da pronúncia do inglês no Brasil e nos demais países sul-americanos.

## 0. INTRODUCTION

With the importance of English as an international language, some special attention to pronunciation should be given in English as a foreign language (EFL) classes, since the reconceptualization of the role of English in the world has generated different purposes for learning it. Seidlhofer (2001) discusses the importance of stress and the amount of effort that should be put into the teaching of significantly different sounds between the native language (NL) and the target language (TL) compared to easier sounds that could be less emphasized. In this sense, Seidlhofer (2001) claims that teachers should be aware of «how much *work* individual sounds, or sound contrasts, actually do in a particular language, that is, whether they have a high or low functional load» (p. 59).

Characterized in this way, the aim of this study is to investigate the difficulty in the production of initial /s/ clusters by Portuguese and Spanish EFL speakers. This difficulty has already been investigated with Spanish speakers by Carlisle (1991, 1992, 1997) and with Portuguese speakers by Rebello (1997). In their studies, the production of epenthesis (the addition of an extra vowel) to the initial clusters was the usual strategy for dealing with syllable structure difficulty, given that in both Portuguese and Spanish, /s/-clusters are invariably preceded by a vowel.

Since Carlisle and Rebello report divergent findings, this investigation aimed at verifying and explaining the different results obtained by the two researchers. Following Carlisle's (1991, 1992, 1997) and Rebello's (1997) studies, three variables were investigated as to their influence on the production of /s/ clusters: the structure of the cluster, its length and the phonological context where it occurs.

This article was divided into five sections. The first section presents a brief review of the literature on the syllable structures of

English, Portuguese and Spanish, as well as some studies on IL syllable structure based on the Markedness Differential Hypothesis (MDH). The second section describes the hypotheses elaborated for this study and the method adopted to collect and analyze the data. The third section provides the analysis of the results, which are discussed in the following section. The fifth section discusses some pedagogical implications, limitations and suggestions for further research.

## 1. SYLLABLE STRUCTURE OF THE LANGUAGES IN QUESTIONS

Since the aim of this study is to investigate how Brazilian Portuguese and Argentinean Spanish speakers produce English initial /s/ clusters, the main syllable structure features of each language will be briefly described, and some studies that show the main strategies used by learners to avoid this difficulty will be reported.

According to Brinton (2000, p. 65), the onset in English may optionally contain up to three consonants and the coda may be formed by one to four consonants, which can be represented as (C) (C) (C) (V) (C) (C) (C) (C). This wide variety of English syllable types can be considered one of the reasons EFL learners face difficulty in producing them correctly. Consonant clusters constitute an important aspect of restrictions on syllable types, since there is a limited number of allowed combinations of segments in both initial and final positions. Clusters with initial /s/, for instance, are the only instances of onsets where the second consonant may be an obstruent and where the onset may be formed by three consonants instead of one or two.

When considering the possible combinations of segments within the syllable structures of both Portuguese and Spanish, it is possible to observe that there is a rather limited number of clusters. The phoneme /s/, for instance, occurs in syllable-initial position when not followed by any other consonant in the three languages; however, it forms initial clusters only in English. This difference is significantly relevant for the study on the production of initial /s/ clusters by EFL learners, since it accounts partially for the difficulty in producing this structure.

In order to better understand the findings related to the production of the /s/ clusters, the Markedness Differential Hypothesis (MDH) and the Structural Conformity Hypothesis (SCH), both elaborated by Eckman (1977) and Eckman (1996), were the theories chosen to give support to the investigation. The first theory proposes that language universals and NL transfer predict difficulties in TL learning. This means that TL structures which are different from and more marked than the corresponding NL structures will be difficult to learn and that the relative degree of difficulty “will correspond to the relative degree of markedness”. Katamba (1989, p. 98) defines markedness in terms of *naturalness*: what is *natural* can be considered *unmarked* in a language, and what is *not natural* can be considered *marked* or in some sense *unusual*. While the MDH makes predictions on the basis of both universals and differences between the NL and the TL, the second theory, the SCH, makes predictions only on the basis of universals, that is, it considers the tendency interlanguages (learners’ transitory competence or intermediary system between NL and TL) have to follow the same universal principles that primary languages do.

There are several studies that show the main strategies used by learners to avoid syllable structure difficulty. Some of them, as previously stated, were highly relevant and were considered the starting point for this study. Carlisle carried out several studies involving native Spanish-speaking learners of English as a second language, who were asked to read a number of topically unrelated and randomly ordered sentences containing initial /s/-clusters in two types of environment. Carlisle (1991) examined epenthesis before three word-initial onsets in English: /sp/, /st/ and /sk/; Carlisle (1992) investigated the production of epenthesis before the word-initial onsets /sl/, /sm/ and /sn/; and Carlisle (1997) compared the production of /sC/ vs. /sCC/ clusters. All the studies controlled the environments before the onsets and the sonority relationships among the consonants in the onsets. His findings in the three studies revealed that vowel epenthesis was significantly more frequent after consonants than after vowels. Concerning the variable *length*

examined in Carlisle (1997), his findings reveal that tri-literal clusters, whose structure is more marked, were more frequently modified than bi-literal clusters, confirming that language universals “influence the structuring of interlanguage phonology” (p. 327).

In another study, Rebello (1997) adopted a type of instrumentation similar to that of Carlisle, but investigated Portuguese-speaking learners of English. She included six English two-segment /s/ clusters (/sp, st, sk, sm, sn, sl/), and five three-segment /s/ clusters (/spr, str, skr, spl, skw/) and found results contrary to those of Carlisle (1991, 1992, 1997). Concerning the *phonological environment*, epenthesis occurred most frequently after silence, followed by vowels and finally by consonants. Although there was little difference in the results concerning the variable *length*, Rebello’s (1997) findings are contrary to those of Carlisle (1997), since epenthesis was more frequently produced before bi-literal clusters than tri-literal clusters, the latter being more marked. As to the structure of the word-initial onsets, Rebello found that bi-literal clusters not in violation of Hooper’s (1976) sonority hierarchy principle were more frequently modified than bi-literal clusters in violation and than tri-literal clusters, which are all in violation. Her results contradict expectations based on Hooper (1976, p. 206) that clusters in violation should be more frequently modified than clusters not in violation. Rebello’s (1997) explanation for these unexpected findings regarding *length* and *structure* of the cluster is that participants tended to voice the /s/ before /s/-nasal and /s/-liquid clusters (transfer of the voicing assimilation process in Portuguese), which resulted in *voiced obstruent + sonorant* clusters. Since voiced obstruents are more marked than voiceless obstruents in any position, these clusters can be considered to be more marked than *voiceless obstruent + obstruent* clusters. Thus, she considers her results to conform to the MDH and the SCH, since “the more marked type of cluster caused more epenthesis than the less marked type”. This difference in results deserved much attention in the present study in order to confirm whether the NL does have such an influence on TL production as to result in opposite findings.

## 2. METHOD

*Participants:* nine native Spanish speakers from Argentina, enrolled in the introductory, first and second years of the *Letras* course at the University of Montoya, in the city of Posadas, in the province of Misiones; and ten native Portuguese speakers from Brazil, enrolled in the second and third years of the *Letras* course at the Universidade Federal de Santa Catarina (UFSC), in the city of Florianópolis.

*Material:* participants were asked to read 180 topically-unrelated sentences, an instrumentation similar to that in both Carlisle's and Rebello's studies. This technique was preferred to guarantee all relevant phonological contexts were present in the corpus. The corpus included 13 sentences containing one occurrence of each of the bi-literal and tri-literal /s/ clusters: /sp, st, sk, sw, sm, sn, sl, spr, str, skr, spl, skw/, each preceded by 5 vowels, 5 consonants, and 3 null contexts (silence). This gave a total of 156 target sentences, to which were added 24 distractor sentences, meaning each participant read 180 sentences. Each subject read the sentences in a different order to prevent any possible ordering effect. Some of the sentences used in this study were taken with permission from Rebello's (1997) corpus.

*Transcription:* only the part of each sentence considered relevant to this study was transcribed. Three aspects were focused on: The absence or presence of the epenthetic vowel, the phonetic realization of the preceding environments, and the phonetic representation of the onsets. The relevant sections of each sentence were first transcribed by the researcher twice, within the interval of a week. Then the items were independently transcribed by one more judge with experience in phonetic transcription. Consequently, out of the 2,964 items read, 2,808 (94.74%) were left for statistical analysis.

*Statistical analysis:* the chi-square ( $\chi^2$ ) was used to analyze the variables considered to be possible constraints in the production of accurate segments. This procedure allows the testing of significance concerning the association of qualitative variables (BARBETTA, 2001). According to Woods, Fletcher and Hughes (1986, p. 144), in order for the chi-square to have satisfactory properties, "all expected

frequencies must be sufficiently large”, and this was the case in this study, since there were a great number of items.

*Hypotheses:* based on the studies by Carlisle and Rebello, the following hypotheses were formulated concerning, length, structure and phonological context where epenthesis most frequently occurs: (1) Spanish speakers would tend to modify longer clusters, while Portuguese speakers would modify shorter clusters more frequently; (2) Spanish speakers would tend to modify more marked structures, while Portuguese speakers would modify less marked structures more frequently; (3) Spanish speakers would tend to produce more epenthesis after consonants, followed by vowels, and then by silence, while Portuguese speakers would epenthesize the sibilant more frequently after silence, followed by consonants and then by vowels.

### 3. RESULTS

#### 3.1. General analysis of /sC/ vs. /sCC/

The results of the first analysis are shown in Table 1 and reveal that both Brazilian and Argentinean participants produced significantly more epenthesis before longer clusters, since there was a 9.62% average difference between /sC/ and /sCC/ clusters within the Brazilian group ( $X^2(1, N = 1,493) = 15.19, p < .0001$ ), and a 10.72% difference within the group of Argentineans ( $X^2(1, N = 1,315) = 17.98, p < .0001$ ).

Table 1 – Total rates of epenthesis production for /sC/ and /sCC/ clusters by Portuguese and Spanish speakers

|    | /sC/          |              |              | /sCC/         |              |              |
|----|---------------|--------------|--------------|---------------|--------------|--------------|
|    | # productions | # epenthesis | % epenthesis | # productions | # epenthesis | % epenthesis |
| BR | 866           | 251          | 28.98        | 627           | 242          | 38.60        |
| AR | 774           | 186          | 24.03        | 541           | 188          | 34.75        |

### 3.1.1 Analysis of clusters grouped by second component

In order to analyze the length of clusters without the intervening variable *sonority*, the bi-literal and tri-literal clusters were compared grouped by their second component, as shown in Table 2.

Table 2 – Total rates of epenthesis production for clusters grouped by their second component

|       | /sp/ |      |              | /spr/ |      |              | /spC/ |      |              | Both /spC/ |      |              |
|-------|------|------|--------------|-------|------|--------------|-------|------|--------------|------------|------|--------------|
|       | Prod | Epen | Rate         | Prod  | Epen | Rate         | Prod  | Epen | Rate         | Prod       | Epen | Rate         |
| BR    | 125  | 33   | <b>26.40</b> | 125   | 45   | <b>36.0</b>  | 123   | 49   | <b>39.84</b> | 248        | 94   | <b>37.90</b> |
| AR    | 114  | 30   | <b>26.32</b> | 112   | 36   | <b>32.14</b> | 99    | 27   | <b>27.27</b> | 211        | 63   | <b>30.81</b> |
|       | /st/ |      |              | /str/ |      |              |       |      |              |            |      |              |
| BR    | 129  | 47   | <b>36.43</b> | 129   | 50   | <b>38.76</b> |       |      |              |            |      |              |
| AR    | 110  | 41   | <b>37.27</b> | 113   | 49   | <b>43.36</b> |       |      |              |            |      |              |
|       | /sk/ |      |              | /skr/ |      |              | /skw/ |      |              | Both /skC/ |      |              |
| BR    | 115  | 34   | <b>29.57</b> | 126   | 47   | <b>37.30</b> | 124   | 51   | <b>41.13</b> | 250        | 98   | <b>39.20</b> |
| Group | 109  | 29   | <b>26.61</b> | 108   | 32   | <b>29.63</b> | 110   | 45   | <b>40.91</b> | 218        | 77   | <b>35.32</b> |

As can be observed, although the tendency in the analysis of clusters grouped by second component was to conform to the MDH, since the more marked clusters were more frequently modified, the results are inconclusive, given that difference in epenthesis production between the total rates were not statistically significant, with the only exception being the analysis of /sp/ versus /spC/ clusters produced by Portuguese speakers, which resulted in a significant chi-square ( $X^2(1, N = 373) = 4.39, p < .04$ ) and confirmed that longer clusters are more difficult to produce, since they resulted in a greater rate of epenthesis.

### 3.2. Analysis of bi-literal clusters in violation of the Syllable Structure Condition (SSC) vs. bi-literal clusters not in violation

Another analysis was carried out in order to investigate whether strength relations within the clusters exerted some influence on the

production of epenthesis. In this analysis, bi-literal /s/ clusters in violation of the sonority hierarchy principle within the syllable (/sp/, /st/, and /sk/) elaborated by Hooper (1976) were compared to bi-literal /s/ clusters not in violation (/sw/, /sm/, /sn/, /sl/), thus eliminating the intervening variable *length*. In Rebello (1997), the findings did not support the hypothesis based on the SSC, because her participants produced more epenthesis before bi-literal clusters not in violation than bi-literal clusters in violation.

However, the results obtained in this study seem to contradict Rebello's (1997) findings, corroborating the SSC. The figures described in Table 3 show the total frequencies of epenthesis produced by Brazilian and Argentinean participants in the analysis of the structure of the clusters. The total rate of epenthesis produced by these groups revealed that a vowel was more frequently inserted before bi-literal clusters in violation than before bi-literal clusters not in violation, although the figures resulted in a non-significant chi-square for Brazilians ( $X^2(1, N = 866) = 1.31, p > .25$ ), and a highly significant chi-square for Argentineans ( $X^2(2, N = 774) = 12.25, p < .0005$ ).

Table 3 – Total rates of epenthesis production for clusters not in violation of the SSC vs. clusters in violation

| Participants | Bi-literal clusters not in violation |              |              | Bi-literal clusters in violation |              |              |
|--------------|--------------------------------------|--------------|--------------|----------------------------------|--------------|--------------|
|              | # productions                        | # epenthesis | % epenthesis | # productions                    | # epenthesis | % epenthesis |
| BR           | 497                                  | 137          | <b>27.57</b> | 369                              | 144          | <b>30.89</b> |
| AR           | 441                                  | 86           | <b>19.50</b> | 333                              | 100          | <b>30.03</b> |

### 3.2.1. Voicing assimilation of /s/-sonorant clusters by Portuguese speakers

Voicing assimilation before sonorants does seem to influence the production of epenthesis by Portuguese speakers. The total frequency of voicing assimilation and the rate of epenthesis of /sN/ and /sl/ clusters is summarized in Table 4. As can be observed, participants voiced the /s/ in 55.60% of /sN/ clusters, and out of these, 56.83% were pronounced with epenthesis. A much lower rate

of epenthesis, 11.71%, was obtained in the productions of clusters where the sibilant was not voiced. Still concerning the category of /sN/ clusters, the contrast between the production of epenthesis in voiced items and in voiceless items resulted in a very significant chi-square ( $X^2(1, N = 250) = 52.10, p < .0001$ ). Similar rates were obtained in the production of the /sl/ cluster. In this cluster, the sibilant was voiced in 59.35% of the items, and out of these, 53.42% were pronounced with epenthesis. Again, epenthesis production without the voicing of the sibilant was considerably lower, 12.0%. The difference between the percentages obtained for epenthesis in voiced and voiceless /sl/ clusters also resulted in a very significant chi-square ( $X^2(1, N = 123) = 20.19, p < .0001$ ).

Table 4: Rates of epenthesis and voicing assimilation by Portuguese speakers

|                       | /s/-nasal     | /s/-liquid    |
|-----------------------|---------------|---------------|
| Number of [+vd] items | 139 (55.60%)  | 73 (59.35%)   |
| Number of epenthesis  | 79            | 39            |
| Rate of epenthesis    | <b>56.83%</b> | <b>53.42%</b> |
| Number of [-vd] items | 111 (44.40%)  | 50 (40.65%)   |
| Number of epenthesis  | 13            | 06            |
| Rate of epenthesis    | <b>11.71%</b> | <b>12.0%</b>  |

### 3.3. The influence of phonological environment in the production of epenthesis

The results of this study partially corroborate those of Rebello (1997), in that epenthesis was more frequently produced after a word-final vowel than after a word-final consonant. However, whereas Rebello (1997) found the greatest frequency of epenthesis at the beginning of an utterance (the null context), in this study, silence was the context which yielded the lowest frequency of epenthesis production. Considering the whole group of Brazilian participants, Table 5 shows that epenthesis occurred more frequently

after open syllables, 40.70%, followed by closed syllables, 32.12%, and by silence, 21.80%. These figures resulted in a very significant chi-square ( $X^2(2, N = 1,493) = 34.98, p < .0001$ ). The results obtained with Argentinean participants, however, corroborate those of Carlisle (1991, 1992, 1997), since all the participants produced more epenthesis after a word-final consonant, 39.64%, followed by a word-final vowel, 22.65%, and then by silence, 16.88%, which resulted in a significant chi-square ( $X^2(2, N = 1,315) = 60.48, p < .0001$ ).

Table 5 – Rates of epenthesis production by Portuguese and Spanish speakers in different environments

|           | Null       |           |         | Vowels     |            |         | Consonants |            |         |
|-----------|------------|-----------|---------|------------|------------|---------|------------|------------|---------|
|           | # prod.    | # epen.   | % epen. | # prod.    | # epen.    | % epen. | # prod.    | # epen.    | % epen. |
| <b>BR</b> | <b>344</b> | <b>75</b> | 21.80   | <b>570</b> | <b>232</b> | 40.70   | <b>579</b> | <b>186</b> | 32.12   |
| <b>AR</b> | <b>314</b> | <b>53</b> | 16.88   | <b>499</b> | <b>113</b> | 22.65   | <b>502</b> | <b>199</b> | 39.64   |

#### 4. GENERAL DISCUSSION

Results concerning the first hypothesis, that longer clusters would be more difficult to produce for Spanish speakers and shorter clusters for Portuguese speakers, reveal that both Portuguese and Spanish EFL speakers inserted an epenthetic vowel more frequently after tri-literal than after bi-literal clusters, confirming the predictions based on the MDH, which claims that more marked structures are more frequently modified. The hypothesis for Spanish speakers was thus confirmed, but not for Portuguese speakers. This might be explained by the fact that in Rebello (1997) differences between shorter and longer clusters were extremely small, since her participants did not read the sentences fluently, given that they had a low level of English proficiency, which led them to produce epenthesis so frequently that the length of the clusters did not seem to be an influencing variable.

The second hypothesis claimed that clusters in violation of the SSC would be more frequently modified by Spanish speakers than

clusters not in violation, and that clusters not in violation would be more frequently modified by the Portuguese speakers than clusters in violation. This hypothesis was confirmed for the Spanish speakers, corroborating the results found in previous studies by Carlisle (1991, 1992, 1997). The findings regarding the Portuguese speakers were inconclusive, which probably indicates the neutralization of opposing tendencies generated by the two different kinds of markedness: The first referring to the SSC, and the second referring to voicing.

The third hypothesis, which aimed at determining whether the phonological environment could be considered a variable constraint, predicted that Portuguese EFL speakers would produce a higher frequency of epenthesis after silence, followed by vowels and then by consonants, and that Spanish EFL speakers would insert an epenthetic vowel more frequently after word-final consonants, followed by vowels and then by silence. The results for the Spanish speakers followed the expected order. However, the results concerning the Portuguese speakers only partially confirm the hypothesis, since epenthesis was more frequently produced after vowels than after consonants, but silence was the context that produced the least rather than the most frequent epenthesis. This means that strength relations across the syllable affected the difficulty in producing English initial /sC(C)/ clusters, but differently for the two NL groups.

Therefore, some results found in this study are contrary to Rebello's (1997) findings. The production of Portuguese EFL speakers was analyzed in both studies; however, the difference in results might be explained by the participants' different levels of English proficiency. While in Rebello's investigation participants were learners attending English classes at the extracurricular course of a public Brazilian university, in the present study participants were learners attending the *Letras* course of the same university, which means that they were more proficient and already had some knowledge of English phonology, given that they are future EFL teachers. Thus, although the findings in this study are different from Rebello (1997), the results show that the production of /s/ clusters by Portuguese speakers is very different from that of Spanish speakers.

Concerning the production of /s/ clusters by Spanish speakers, the results corroborate Carlisle's (1991, 1992, 1997) findings, since (1) participants modified longer clusters more frequently than shorter cluster; (2) the insertion of an epenthetic vowel occurred more frequently before clusters in violation of the sonority hierarchy than before clusters not in violation; and (3) epenthesis occurred more frequently after vowels, followed by consonants, and then by silence. Considering, then, what was found by Carlisle and Rebello, the difference in results does not seem to have been caused by the differences in corpora. Differences in Spanish and Portuguese L1 show that native language interference and language universals can more accurately explain the dissimilarities.

## 5. PEDAGOGICAL IMPLICATIONS

The results of this and other studies on interlanguage phonology should convince materials writers to design pronunciation materials by presenting less marked TL structures before more marked structures. In this sense, bi-literal clusters should be presented before tri-literal clusters, clusters not in violation should be presented before clusters in violation, at least for Spanish speakers, and the /sl/ cluster should be presented before /sN/ clusters. Concerning phonological environment, this and other studies investigating Spanish EFL speakers found that /s/ clusters in vocalic environments cause less difficulty; thus they should be presented first to these learners. As regards Portuguese EFL speakers, this study and Rebello (1997) lead to the suggestion that consonantal and voiceless environments should be presented before vocalic and voiced environments.

Thus, research in interlanguage phonology would play an important role in helping both teachers and learners if the findings obtained in this interphonology research were taken into consideration in the development of language-specific series of English pronunciation manuals or textbooks that focus on the difficulties of each native language learner group.

One limitation of this study has to do with the participants' level of proficiency, since they had already received phonological instruction regarding /s/ clusters, which is an aspect that influences the results of epenthesis production. Thus, an analysis encompassing EFL beginners and fluent speakers could reveal to what extent the variable *proficiency* plays a role in influencing the production of epenthesis before English initial /s/ clusters.

In addition to the inclusion of several proficiency levels and both elicited and natural speech, the difference between acquisition in natural settings and in the language classroom could also be analyzed. It is worth stressing that the studies by Carlisle were carried out in the United States, where learners were acquiring the language in a natural environment, unlike the learners in this study. However, a systematic study comparing the two environments would be useful, since it could contribute to a better understanding of how the learning process takes place in an informal versus a classroom environment.

## REFERENCES

- BARBETTA, P. A. *Estatística aplicada às ciências sociais*. Florianópolis: Editora da UFSC, 2001.
- BRINTON, L. J. *The structure of modern English: A linguistic introduction*. Amsterdam/ Philadelphia: John Benjamins Publishing Company, 2000.
- CARLISLE, R. S. The influence of environment on vowel epenthesis in Spanish/English interphonology. *Applied linguistics*, v.1, n. 12, p. 76-95, 1991.
- CARLISLE, R. S. Environment and markedness as interacting constraints on vowel epenthesis. In: LEATHER, J.; JAMES, A. (Org.). *New Sounds 92*. University of Amsterdam, p. 64-75, 1992.
- CARLISLE, R. S. The modification of onsets in a markedness relationship: Testing the interlanguage structural conformity hypothesis. *Language learning*, v. 2, n. 47, p. 327-361, 1997.
- ECKMAN, F. R. A functional-typological approach to second language acquisition theory. In: RITCHIE, W. C.; BHATIA, T. K. (Org.). *Handbook of second language acquisition*. San Diego: Academic Press, 1996. p. 195-211.

HOOPER, J. *Introduction to natural generative phonology*. New York: Academic Press, 1976.

KATAMBA, F. *An introduction to phonology*. London/NewYork: Longman, 1989.

REBELLO, J. *The acquisition of initial /s/ clusters by Brazilian EFL learners*. Unpublished master's thesis. Universidade Federal de Santa Catarina, 1976.

SCHMIDT, R. The role of consciousness in second language learning. *Applied Linguistics*, n. 11, p. 129-158, 1990.

SEIDLHOFER, B. Pronunciation. In: CARTER, R.; NUNAN, D. (Org.). *The Cambridge guide to teaching English to speakers of other languages*. Cambridge: Cambridge University Press, 2001. p. 56-65.

WOODS, A.; FLETCHER, P.; HUGHES, A. *Statistics in language studies*. Cambridge, Cambridge University Press, 1986.

# Categorias fonéticas na aprendizagem de língua estrangeira

Barbara O. Baptista  
Universidade Federal de Santa Catarina

## Abstract

The purpose of this paper is, through a brief revision of cross-language research on the phonetic perception and production of native and foreign/second languages, to propose an extension, in two directions, to Flege's (1995) *Speech Learning Model (SLM)* and to discuss the implications for Best's (1995) *Perceptual Assimilation Model (PAM)* of a broader view of *phonetic category*. The two directions for an extension of the *SLM* are (a) to include, among the phonetic categories to be learned/acquired, units larger than the segment, such as the syllable, onset, or coda; and (b) to incorporate the idea that many of the segments are acquired as part of a natural class rather than in isolation. The discussion of the *PAM* deals with the effect that the consideration of larger phonetic categories and natural classes would have on Best's proposed patterns of assimilation.

## 1. INTRODUÇÃO

**N**a pesquisa sobre a aquisição ou aprendizagem dos sons de uma língua estrangeira (LE) da última década, os dois modelos teóricos que têm exercido a maior influência são, sem dúvida, o *Speech Learning Model (SLM)* – Modelo de Aprendizagem da Fala – de Flege (1995) e o *Perceptual Assimilation Model (PAM)* – Modelo de Assimilação Perceptiva – de Best (1995). O *SLM*, com seus quatro postulados e sete hipóteses, sustenta que o adulto mantém a capacidade de percepção fonética da infância e consegue formar novas categorias fonéticas para os sons claramente distintos da LE. A dificuldade seria com os sons da LE semelhantes aos da língua materna (LM), que seriam percebidos como pertencendo a categorias da LM, devido a um mecanismo denominado *classificação por equivalência*. O *PAM* propõe três padrões de assimilação perceptiva dos sons da LE às categorias fonéticas da LM e seis padrões de assimilação de pares de sons contrastantes da LE, com várias combinações de assimilação a uma, duas ou nenhuma categoria da LM. A cada padrão de assimilação de pares contrastantes é atribuído seu esperado grau de discriminação.

Apesar da discussão nas últimas décadas sobre a unidade de representação mental dos sons da língua falada, os dois modelos se referem apenas à unidade segmental isoladamente, como se o segmento fosse a única possibilidade de representação e como se os segmentos fossem adquiridos um a um, sem referência uns aos outros. Essa limitação claramente não foi por descuido, mas por uma decisão consciente de ambos os autores.

O primeiro postulado do *SLM* parece reconhecer a importância do *sistema* fonético, pelo menos da LM, quando afirma que “os mecanismos e processos usados na aprendizagem do *sistema de sons* [itálico meu] da LM,..., permanecem intactos ao longo da vida”

(p. 239). Entretanto, os outros postulados e todas as hipóteses referem-se aos sons individuais, sem menção de ligação entre eles. Quanto à unidade de representação, Flege esclarece de início que “embora as dimensões não segmentais (i.e., prosódicas) sejam uma importante fonte de sotaque estrangeiro, o presente capítulo focaliza em unidades da fala nas dimensões de fonemas” (p. 233). Mais adiante, o autor nega, de maneira explícita, a aplicação do *SLM* a unidades maiores de representação, atribuindo dificuldades com seqüências da LE a limitações motoras:

O modelo *não* afirma, contudo, que todos os erros de produção de LE sejam atribuídos à percepção. Por exemplo, limites na produção motora baseados em tipos de sílaba permitidos da LM podem fazer com que falantes do espanhol pronunciem a palavra ‘school’ como [ɛskul]. (1995, p. 238)

Best (1995) parece reconhecer uma certa ligação entre as unidades fonéticas quando se refere às “partes do domínio fonético que uma língua utiliza para realizar as funções lingüísticas” (p. 186), que ela chama de “espaço fonológico nativo” (p. 187). No que diz respeito à representação mental, Best opina que esta é gestual e reconhece unidades maiores do que o fonema quando se refere às “constelações gestuais” de Browman e Goldstein (1990a, 1990b), que ela descreve assim:

As línguas diferem de maneira dramática na sua junção de gestos simples em constelações gestuais. Tais estruturas fonológicas certamente não estão limitadas apenas ao nível segmental de descrição lingüística, mas estendem-se a outros níveis fonológicos, inclusive às sílabas e seus constituintes, às unidades rítmicas, etc. (p. 192)

Apesar dessa concessão, a autora resolve focalizar nos segmentos e contrastes segmentais, justificando-se pela pesquisa existente sobre a percepção dos sons da LE, a grande maioria limitada a contrastes e comparações segmentais. Best faz referência a unidades maiores apenas mais uma vez – na descrição do primeiro padrão de assimilação perceptiva: “claramente assimilado a uma determinada categoria segmental nativa, ou talvez a um grupo ou seqüência” (p. 194).

Embora esses modelos tenham contribuído muito para a compreensão do conhecimento fonético do falante de uma língua estrangeira, proponho que é chegada a hora de se reconhecer a provável multiplicidade de unidades de representação fonética, bem como a necessária ligação entre uma categoria fonética e a outra, e de se começar a estender os modelos teóricos de maneira que sejam levadas em consideração as outras unidades de representação e as classes fonológicas naturais.

## 2. A UNIDADE DE REPRESENTAÇÃO FONÉTICA

Como mencionado na introdução, existe, há pelo menos duas décadas, uma polêmica sobre a unidade de representação mental dos sons, seja na LM ou na LE, que tem motivado muitas pesquisas experimentais nas áreas da percepção lingüística e do acesso lexical. Há vários trabalhos que fornecem evidências para a sílaba ou uma seqüência menor, como o ataque, como unidade de representação; e há trabalhos que sugerem diferenças na unidade perceptiva, dependendo da língua dos falantes.

Cutler, Butterfield e Williams (1987) mediram o tempo de detecção por falantes nativos de consoantes iniciais em ataques simples e em ataques complexos (encontros consonantais) em palavras inglesas contextualizadas. Nas palavras de ataques complexos, o tempo de detecção variava em função do ataque da palavra alvo conter ou não a mesma segunda consoante que o ataque do modelo. Nas palavras de ataques simples, o tempo não dependia da vogal seguinte. Baseados nos resultados, os autores sugerem que encontros consonantais iniciais são percebidos como unidades integrais.

Levelt e Wheeldon (1994) realizaram quatro experimentos em que falantes nativos de holandês aprenderam a associar cada uma de 16 palavras bissilábicas holandesas (palavras diferentes em cada experimento) a um conjunto de símbolos abstratos. Constataram que o tempo para os sujeitos nomearem cada palavra, quando apresentados ao conjunto de sílabas, era afetado pela frequência da segunda sílaba,

independente do efeito da frequência da palavra e independente da complexidade da sílaba. Os resultados apoiam a hipótese de que a maioria das sílabas produzidas por um falante são gestos articulatórios altamente automatizados, armazenados e acessados como unidades prontas.

Cutler, Mehler, Norris e Segui (1986) pediram para falantes de francês e de inglês realizarem tarefas de monitoramento de seqüências CV e CVC em listas de palavras. O menor tempo de detecção, somente pelos franceses, das seqüências com a mesma estrutura silábica que o modelo levaram os autores à conclusão de que a estratégia de compreensão desenvolvida por um indivíduo depende da estrutura da sua língua: o francês incentiva a estratégia silábica, mas o inglês não. Assim podemos inferir que a sílaba constitui uma representação fonética importante para os falantes do francês, mas não para os falantes do inglês.

Essas são apenas três de muitas pesquisas realizadas com o objetivo de descobrir a unidade básica de representação e processamento lingüístico. Talvez a conclusão mais sensata seja a de Pisoni e Luce (1986, 1987) de que há uma variedade de unidades usadas no processamento da fala. Essa atitude não se reflete na pesquisa empírica sobre a percepção e produção de língua estrangeira, nem nas principais teorias, apesar de algumas asserções terem implicações nesse sentido.

No *SLM* a unidade de representação é o alofone posicional. Flege (1995) cita vários trabalhos que fornecem evidências de que o aprendiz de língua estrangeira percebe os sons de maneira diferente dependendo da posição deles. Strange (1992, *apud* FLEGE, 1995, p. 239) sustenta que aprendizes japoneses de inglês geralmente percebem e produzem as líquidas do inglês melhor em posição final de palavra do que em posição inicial. De maneira semelhante, Takagi (1993, *apud* FLEGE, 1995, p. 239), em uma experiência em que participantes japoneses tinham que escrever os símbolos katakana para as sílabas do inglês que ouviam, demonstrou que ambas as sílabas inglesas /lV/ e /rV/ eram percebidas como /rV/, mas que o /r/ final era percebido geralmente como /a/ e o /l/ final como /ru/ ou /o/.

Essas pesquisas realmente apoiam a hipótese de que os “alvos perceptuais” sejam posicionais, mas levam também a uma implicação maior: quando um segmento – o /l/ no trabalho de Takagi – é percebido como uma sílaba – /ru/ – isso não é evidência da existência dessa sílaba como categoria fonética? Enfim, é difícil um alofone estar ligado a uma posição na representação mental sem haver uma representação também para essa posição. Desse modo, a conclusão lógica é de que a representação mental fonética consiste em categorias fonéticas abstratas para os sons individuais, cujos alofones se encontram dentro de categorias fonéticas maiores para as seqüências contextuais, que podem ser grupos de segmentos ou sílabas.

A primeira hipótese do *SLM* sustenta que “os sons da LM e da LE estão relacionados perceptivamente em nível alofônico posicional ao invés de em nível fonêmico abstrato” (FLEGE, 1995, p. 238). O autor prevê, baseado nessa hipótese, que falantes de uma LM que não tem oclusivas finais não relacionarão as oclusivas finais do inglês, por exemplo, às oclusivas mediais ou iniciais e que, assim, sendo sons novos, os aprendizes chegarão a produzir essas oclusivas de maneira adequada (p. 261). Entretanto, ele cita vários trabalhos em que os aprendizes omitem ou desvozeiam as oclusivas finais ou acrescentam uma vogal epentética, e discute várias explicações possíveis pelos resultados contraditórios.

Dois trabalhos recentes também são relevantes para essa discussão. Baptista e Silva (1997) e Koerich (2002) investigaram a produção de consoantes finais do inglês por brasileiros e constataram que a estratégia mais comum foi o acréscimo de vogal epentética. Algumas constatações de Baptista e Silva – e.g., o maior índice de epêntese após as oclusivas velares vozeadas do que após as alveolares e bilabiais, devido a diferenças na pressão supraglotal – apontam para uma explicação articulatória. No entanto, o resultado mais importante de Koerich – a correlação entre a produção de epêntese e a dificuldade de percepção da distinção entre seqüências CVCV e seqüências CVC – aponta para uma explicação em termos perceptuais. Não é difícil sustentar a hipótese (H1) de que o aprendiz de LE relacione

os sons da LM e da LE em nível alofônico posicional, quando as estruturas silábicas das duas línguas são semelhantes. Mas quando a LE permite consoantes finais e a LM não, além de não haver, na representação mental, alofone posicional na LM com que fazer a relação, não há nem posição em que criar o alofone novo. Deste modo, o aprendiz só pode fazer a relação com o alofone posicional que existe na LM, no caso, o alofone medial. Assim, ele nem percebe a ausência da vogal final no insumo da LE e ouve duas sílabas CV, para as quais ele tem uma categoria fonética, ao invés de uma sílaba CVC, porque, para esta sílaba, falta-lhe a categoria fonética.

Outra seqüência problemática do inglês, objeto de pesquisas recentes com falantes nativos de espanhol e português, são os encontros consonantais iniciais com /s/: /sp, st, sk, spr, spl, str, skr, skw, sm, sn, sl/ (CARLISLE, 1992, 1994, 1997, 1998; REBELLO, 1997, RAUBER, 2002). Embora nenhuma dessas pesquisas lidou com percepção, há evidências do ensino da pronúncia do inglês para alunos intermediários brasileiros: vários alunos meus apresentaram grandes dificuldades em exercícios de distinção entre pares como *specially/especially* e *state/estate*. Eu diria, informalmente, que estes eram os mesmos alunos que tinham dificuldades na pronúncia dessas seqüências, mas essa hipótese precisa ainda ser pesquisada de maneira sistemática para que seja possível propor uma categoria fonética para essas seqüências também.

Resumindo, embora os trabalhos existentes sobre a produção, e especialmente sobre a percepção, de seqüências fonológicas por falantes nativos e não nativos sejam bem menos numerosos do que sobre segmentos, os resultados sugerem a possibilidade de categorias fonéticas maiores do que o segmento. A inclusão de unidades maiores em modelos de aprendizagem, percepção e produção de LE seria um incentivo para mais pesquisas nessa linha, o que certamente levaria a avanços na compreensão do conhecimento fonético do falante de uma LE.

### 3. SUPERCATEGORIAS FONÉTICAS E CLASSES NATURAIS

Se há pouca pesquisa sobre unidades maiores que o segmento na área da percepção e produção dos sons de LE, há ainda menos pesquisas que façam referência a classes naturais ou à formação de supercategorias fonéticas. Entretanto, embora os postulados e hipóteses do *SLM* se refiram somente a sons individuais, vários dos trabalhos sobre a percepção e produção das consoantes de LE citados por Flege (1995) como apoio para o modelo deixam claro, se não explícito, que as categorias fonéticas consonantais são subcategorias de categorias maiores, que correspondem a classes naturais.

Flege (1987), por exemplo, em uma investigação de bilíngües inglês/francês e francês/inglês, examinou a influência dos valores de VOT da LE na produção de /p, t, k/ iniciais da LM. De maneira similar, Flege (1991) se refere aos valores intermediários do VOT de /p, t, k/ do inglês como produzidos por falantes do espanhol. Ambos os trabalhos fazem referência implícita à classe natural de oclusivas não vozeadas iniciais. Flege, Munro e MacKay (1995b, *apud* FLEGE, 1995, p. 259-262) investigaram o efeito da idade de aprendizagem do inglês por italianos no seu VOT de /p, t, k/ iniciais e na sua produção e identificação de /p, t, k/ e /b, d, g/ finais do inglês. Aqui também há referência implícita a três classes naturais de consoantes, cujas categorias fonéticas da LM e talvez da LE devem ser armazenadas em representações na memória de longo prazo dentro de três supercategorias.

As vogais da LE, nos estudos citados por Flege (1995), são geralmente investigadas como sons individuais, cujas categorias fonéticas seriam formadas isoladamente. O autor (p. 242) cita Liljencrants and Lindblom (1972) e Lindblom (1990) que apontam para a importância da dispersão vocálica para a manutenção de contraste auditivo suficiente. No entanto, enquanto Lindblom refere-se ao posicionamento relativo de vogais dentro de uma única língua, Flege relaciona a dispersão ao distanciamento de uma vogal da LE de uma vogal da LM. O *SLM* e os trabalhos citados em Flege (1995) fazem referência às vogais adjacentes somente para dizer que uma distinção entre duas vogais está ou não está sendo feita ou para dizer se uma vogal está sendo produzida de maneira mais precisa que outra.

A importância da dispersão das vogais dentro do espaço acústico de uma língua foi demonstrada por Ladefoged e Broadbent (1957), com estímulos sintéticos, e por Ladefoged (1989), com fala natural: os dois trabalhos fornecem evidências de que o ouvinte percebe cada vogal em relação ao espaço vocálico acústico total do falante, que ele calibra a partir dos padrões de frequência dos formantes no resto da fala. A implicação para o aprendiz de LE é que, para que suas vogais sejam bem percebidas, elas têm que ser produzidas com uma distribuição adequada dentro desse espaço. Para isso, suas categorias fonéticas vocálicas precisam incluir informações sobre o posicionamento relativo entre elas, o que é possível somente se houver uma ligação entre elas; ou seja, mais uma vez precisamos de uma supercategoria do sistema vocálico, talvez com subcategorias das anteriores, posteriores, etc. Se esse for o caso, a pesquisa que trata as vogais isoladamente pode estar negligenciando informações importantes.

A melhor maneira de se obter um quadro mais satisfatório das influências intralingüísticas entre as vogais da LE é investigar porções maiores dos sistemas vocálicos dos aprendizes, mas até hoje, há poucos trabalhos que fazem isso. Baptista (2000, 2002), em um estudo longitudinal e acústico da aquisição das vogais do inglês por onze brasileiros residindo há pouco tempo nos Estados Unidos, fornece três evidências para a hipótese de que as vogais da interlíngua constituem um sistema tão integrado quanto as vogais da LM: (1) nas fases iniciais da aquisição, as vogais anteriores mantiveram o posicionamento relativo aproximado entre si das vogais da LM, mas todas elas se encontravam deslocadas para a frente; (2) a vogal nova /ʌ/, nas fases iniciais, encontrava-se posicionada corretamente em relação às duas vogais baixas /æ/ e /ɑ/, mas todas três estavam altas demais em termos absolutos; (3) a aquisição das vogais novas /ɪ/ e /æ/ ocorreu em função de ajustes nas vogais adjacentes, similares a vogais da LM.

Da mesma forma, Morais (1995), em um estudo da aquisição das vogais do francês por alunos brasileiros, afirma que, quando os aprendizes brasileiros aprendem a produzir a vogal francesa /y/ mais

fechada, anteriorizada e labializada, eles também produzem um /i/ mais anterior e um /u/ mais posterior, adequando assim a distância perceptiva entre as três vogais altas.

Os trabalhos resumidos aqui fornecem ampla evidência de que na aquisição de uma LE, as categorias fonéticas dos sons não são construídas isoladamente, mas como parte de categorias fonéticas maiores. Assim podemos falar da aprendizagem do *sistema* fonético da LE, com seus vários subsistemas, cada som individual pertencendo a um ou mais deles.

#### 4. UMA EXTENSÃO AO SLM

Essa revisão de pesquisas sobre a percepção e produção de sons na LM e na LE demonstrou a importância da inclusão, no modelo, (a) de unidades de representação fonética maiores do que sons individuais, uma vez que há evidências da percepção (adequada ou não), pelo falante de LE, de seqüências de segmentos como unidades; e (b) do conceito de sistema fonético onde os alofones posicionais se encontrariam dentro de supercategorias fonéticas, que corresponderiam a classes naturais de vários níveis, e.g., o sistema vocálico, as vogais anteriores, os oclusivas vozeadas e não vozeadas. O que segue é uma proposta de uma extensão do *SLM* de Flege, no sentido de incluir esses dois conceitos.

Os postulados e hipóteses do *SLM* ficariam assim, as alterações e os acréscimos indicados em itálico:

##### Postulados

- P1: O ser humano pode aplicar os mecanismos e procedimentos usados para a aprendizagem do sistema de sons da LM durante a vida inteira, inclusive para uma LE. [sem alteração]
- P2: Os aspectos dos sons de uma língua específica são armazenados em categorias fonéticas na memória de longo prazo. *Muitos dos aspectos dos sons seriam relativos aos aspectos de outros (e.g., posicionamento das vogais, VOT de consoantes), o que seria*

*possibilitado pelo armazenamento das categorias fonéticas de sons, com ligações entre si, dentro de supercategorias de classes naturais.*

- P3: As categorias fonéticas construídas na infância evoluem ao longo da vida para refletir as características de todos os sons *e classes de sons* da LM e LE que forem identificados como realizações dos mesmos.
- P4: Os bilingües procuram manter contraste entre os sons, *seqüências de sons e classes de sons* da LM e os da LE, que existem em um espaço fonológico comum.

### Hipóteses

- H1: Os sons da LM e da LE estão ligados perceptivamente em nível alofônico posicional, *por estarem representados dentro de categorias fonéticas de seqüências.*
- H2: O bilingüe pode construir uma nova categoria fonética para um som *ou para uma seqüência de sons* da LE se ele discerne alguma diferença fonética entre este e o som/*seqüência* mais *parecido* da LM.
- H3: Quanto maior a diferença fonética percebida entre um som/*seqüência* da LE e o som/*seqüência* mais *parecido* da LM, maior a probabilidade do bilingüe discernir as diferenças fonéticas. [circular]
- H4: A probabilidade do discernimento entre sons *e seqüências* da LM e LE e entre um par de sons *ou seqüências* da LE que não são contrastivos na LM diminui na medida em que a *idade de aprendizagem* aumenta.
- H5: A formação de categoria para um som/*seqüência* da LE pode ser bloqueada pelo mecanismo de classificação por equivalência. Assim uma única categoria será utilizada para processar os sons/*seqüências* correspondentes das duas línguas, os quais tendem a aproximar-se. [*A aproximação de seqüências provavelmente não acontece.*]
- H6: A categoria fonética construída, por um bilingüe, para um som *ou classe de sons* da LE poderá diferir da categoria de um monolingüe (a) se for desviado de uma categoria da LM ou (b) se for baseado em traços diferentes ou em pesos diferentes para os traços.

H7: A produção de um som/*classe de sons/seqüência de sons* chegará a corresponder às características representadas na categoria fonética.

## 5. IMPLICAÇÕES PARA O PAM

Os mesmos conceitos podem ser incorporados ao *PAM*, nos padrões de assimilação perceptiva de segmentos e nos padrões de assimilação para contrastes dentro da LE, onde é previsto o grau de discriminação para cada. A percepção pode ser de segmentos ou seqüências da LE (não se ouve uma classe natural) e a assimilação pode ser segmentos, seqüências ou classes. Como acima, as modificações são assinaladas em itálico:

Padrões de assimilação perceptiva de segmentos da LE

- (1) Assimilado a uma categoria segmental da LM – ou talvez a um grupo ou seqüência *ou classe*. Ouvido como (a) exemplar bom, (b) exemplar aceitável, (c) exemplar divergente [*Somente neste padrão a autora incluiu seqüências, mas pode-se incluir classes naturais também*];
- (2) Assimilado como som *ou seqüência* da fala, mas não enquadrado em nenhuma *categoria* específica da LM;
- (3) Não assimilado à fala: ouvido como som que não é da fala.

Padrões de assimilação para contrastes dentro da LE

1. Assimilação a duas categorias: cada segmento *ou seqüência* da LE se enquadra em uma categoria diferente: discriminação excelente;
2. Diferença de adequação à categoria: ambos segmentos/*seqüências* se enquadram na mesma categoria, mas com diferentes graus de adequação: discriminação moderada a boa;
3. Assimilação à categoria única: ambos segmentos *ou seqüências* se enquadram na mesma categoria com o mesmo grau de adequação: discriminação ruim.

4. Ambos assimilados como sons *ou seqüências* da fala, mas não enquadrados em nenhuma categoria específica da LM: discriminação de ruim a muito boa, dependendo da distância entre eles e da distância entre cada um deles e as categorias da LM;
5. Categorizado/não categorizado: um segmento *ou seqüência* se enquadra em categoria da LM, o outro dentro do espaço fonético, mas não enquadrado: discriminação muito boa;
6. Não assimilável: não ouvido como som *ou seqüência* da fala: discriminação muito boa.

Além das alterações acima, o modelo ficaria mais completo se fosse acrescentada, às previsões do grau de discriminação dos pares contrastivos, uma previsão do grau de percepção do *relacionamento* entre dois segmentos da LE ou entre um segmento e o resto da classe natural, dentro do espaço acústico. Esse relacionamento seria especialmente importante no caso das vogais, para que cada uma possa ser percebida em relação às outras.

Neste trabalho foi fornecida evidência, a partir de investigações do acesso lexical na LM e da percepção e produção de sons da LE, tanto para a existência de unidades de representação maiores do que o segmento, quanto para a construção de categorias fonéticas interligadas dentro de sistemas fonéticos. Apoiada nessas evidências, propus alterações e acréscimos aos modelos de Flege e de Best. Espero ter contribuído, com essas propostas, para incentivar futuras pesquisas que levem em conta a possibilidade dessas categorias fonéticas maiores e a aprendizagem de sons da LE como parte de supercategorias fonéticas, que abrangem classes naturais de segmentos.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BAPTISTA, B. O. *The acquisition of English vowels by Brazilian-Portuguese speakers*. Florianópolis, SC: Universidade Federal de Santa Catarina, 2000.
- BAPTISTA, B. O. Adult phonetic learning of a second language vowel system. In: LEATHER, J.; JAMES, A. (Org.). *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech (University of Amsterdam 2000)*. Klagenfurt: University of Klagenfurt, p. 32-41, 2002.
- BAPTISTA, B. O., SILVA FILHO, J. L. A. The influence of markedness and syllable contact on the production of English final consonants by EFL learners. In: LEATHER, J., JAMES, A. (Org.). *New Sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech*. Klagenfurt: University of Klagenfurt, p. 26-34, 1997.
- BEST, C. T. A direct realist view of cross-language speech perception. In: STRANGE, W. (Org.). *Speech perception and linguistic experience: Theoretical and methodological issues*. Baltimore, MD: York Press, p. 171-203, 1995.
- CARLISLE, R. S. Environment and markedness as interacting constraints on vowel epenthesis. In: LEATHER, J.; JAMES, A. (Org.). *New Sounds 92: 1992 Amsterdam Symposium on the Acquisition of Second-language Speech*. Amsterdam: University of Amsterdam, p. 64-75, 1992.
- CARLISLE, R. S. Markedness and environment as internal constraints on the variability of interlanguage phonology. In: YAVAS, M. (Org.). *First and second language phonology*, San Diego: Singular, p. 223-249, 1994.
- CARLISLE, R. S. The modification of onsets in a markedness relationship: Testing the Interlanguage Structural Conformity Hypothesis. *Language Learning*, n. 47, p. 327-361, 1997.
- CARLISLE, R. S. The acquisition of onsets in a markedness relationship: A longitudinal study. *Studies in Second Language Acquisition*, n. 20, p. 45-249, 1998.
- CUTLER, A., BUTTERFIELD, S., WILLIAMS, J. N. The perceptual integrity of syllabic onsets. *Journal of Memory and Language*, 26, p. 406-418, 1987.
- CUTLER, A., MEHLER, J., NORRIS, D., SEGUI, J. The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, n. 25, p. 385-400, 1986.
- FLEGE, J. E. The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, n. 15, p. 47-65, 1987.

FLEGE, J. E. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a 2nd language. *Journal of the Acoustical Society of America*, n. 89, p. 395-411, 1991.

FLEGE, J. E. Second language speech learning theory, findings and problems. In: STRANGE, W. (Org.). *Speech perception and linguistic experience: Issues in cross-language research*. Timonium, MD: York Press, 1995, p. 233-277.

KOERICH, R. D. *Perception and production of word-final vowel epenthesis by Brazilian EFL students*. 2001. Tese (Doutorado em Letras/Inglês e Literatura Correspondente) - Universidade Federal de Santa Catarina, Florianópolis.

LEVELT, W. J. M., WHEELDON, L. Do speakers have access to a mental syllabary? *Cognition* n. 50, p. 239-269, 1994.

MORAIS, C. A. *Labialização das vogais orais do sistema vocálico francês por alunos brasileiros: Caso particular /y/, estudo acústico*. 1995. Dissertação (Mestrado em Lingüística) - Universidade Federal de Santa Catarina, Florianópolis.

PISONI, D. B., LUCE, P. A. Speech perception: Research, theory, and the principal issues. In: SCHWAB, E. C., NUSBAUM, H. C. (Org.). *Pattern recognition by humans and machines: Speech perception*. Orlando, FL: Academic Press, 1986. p. 1-50.

PISONI, D. B., LUCE, P. A. Acoustic-phonetic representation in word recognition. In: FRAUENFELDER, U. H.; TYLER, L. K. (Org.). *Spoken word recognition*. Cambridge, MA: The MIT Press, 1987. p. 21-52.

RAUBER, A. S. *The production of English initial /s/ clusters by Portuguese and Spanish EFL speakers*. 2002. Dissertação (Mestrado em Letras/Inglês e Literatura Correspondente) - Universidade Federal de Santa Catarina, Florianópolis.

REBELLO, J. T. The acquisition of English initial /s/ clusters by Brazilian EFL learners. In: LEATHER, J.; JAMES, A. (Org.). *New sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech*. Klagenfurt: University of Klagenfurt, p. 336-350, 1997.

# **A aquisição de encontros consonantais do inglês iniciados por /s/: estudo de caso com criança brasileira**

Magliane De Marco  
Universidade Católica de Pelotas

## **Abstract**

Language acquisition studies have pointed the first language (L1) influence on this process. In a research with adolescents and adults, Fernandes (2001) and Silveira (2002) show that, in the acquisition process of English as foreign language (L2) by Brazilian Portuguese speakers, the L1 influences the L2 syllable structure acquisition. However, concerning to children, it seems that there is less influence of L1 on the foreign language acquisition. Based on that, the aim of this paper is to analyze the word-initial /s/ clusters in English, as well as the possibility of syllable structure transference from first language to L2. The data is from a child that is still developing her phonological system and her age was 3;4 (years/months). She's acquiring Brazilian Portuguese as a first language and English, as a second.

A aquisição fonológica ocorre de forma gradual, do não-marcado para o marcado, o que significa que se adquire primeiro o mais simples para depois, então, adquirir-se o mais complexo. Investigações sobre aquisição da fonologia mostram que a criança possui, inicialmente, a habilidade de produzir todos os sons da fala humana, porém, com o tempo, perde aqueles que não são produzidos em seu ambiente lingüístico (INGRAM, 1989, p. 96). O domínio da fonologia de uma língua implica a aquisição, entre outros aspectos, dos segmentos e das estruturas silábicas que constituem o sistema ao qual a criança está exposta.

Acredita-se que a criança tenha um papel ativo nesse processo de aquisição – o que significa um grau de autonomia –, procurando estabelecer hipóteses e adotando estratégias até alcançar o sistema fonológico do adulto e da comunidade na qual está inserida (MATZENAUER, 1990, p. 1). As primeiras palavras produzidas implicam já a existência de um sistema fonológico, o que não quer dizer que não haja um espaço de tempo entre o entendimento e a produção de palavras (INGRAM, 1989, p. 140). Com o desenvolvimento lingüístico da criança, esse sistema vai sofrendo alterações até chegar ao sistema alvo.

Essa autonomia evidenciada por estratégias ocorre não só na aquisição da língua materna (LM), mas também de uma língua estrangeira (LE). Dessa forma, é interessante que se pesquise a aquisição de uma LE em período em que o sistema da língua materna ainda não está plenamente dominado pela criança, a fim de se verificar, entre outros fatos, se as estratégias aplicadas na aquisição da LM também se fazem presentes no processo de aquisição da LE.

Muitas pessoas aprendem mais de uma língua; no entanto, esse processo pode se dar de várias maneiras. Uma criança pode ser

exposta a duas ou mais línguas desde o início, por exemplo, se seus pais fazem uso de diferentes línguas. Nesse caso, deve-se, ainda, falar em ‘aquisição de língua materna’ – com a diferença de que não somente uma, mas duas línguas são ‘maternas’. Em outras palavras, uma língua é considerada LM se nenhuma outra língua foi adquirida antes; de outra forma é considerada como língua estrangeira. Essa diferença é muito clara se a aquisição da LE começa quando a aquisição da LM já está completa, como no caso de adolescentes e adultos. Se a criança adquire somente uma língua como LM, considera-se ‘aquisição monolíngue de LM’; se a criança está aprendendo duas línguas em paralelo refere-se à ‘aquisição bilíngüe de LM’.

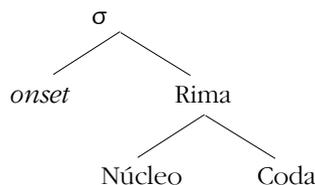
Muitas vezes, decidir entre usar o termo ‘aquisição de LE’ ou ‘aquisição bilíngüe de LM’ é uma questão de preferência pessoal. O uso mais freqüente na literatura refere-se à ‘aquisição de LE’ se ela começa durante ou após a idade de 3 ou 4 anos.

Depois de a sílaba ser aceita como unidade fonológica, os estudos a respeito de sua natureza e do papel desempenhado por ela na fonologia das línguas vêm crescendo cada vez mais.

São várias as teorias que abordam e buscam explicar a estrutura silábica, porém, para este trabalho, será adotado o modelo proposto por Selkirk (1982), uma vez que considera uma estrutura interna para a sílaba, possibilitando dessa forma o estudo de apenas um de seus constituintes.

A sílaba, representada pela letra grega  $\sigma$ , apresenta, portanto, uma estrutura interna: pode ser formada por um ataque ou *onset* e uma rima. A rima pode ser constituída por um núcleo e uma coda como em (1), conforme Selkirk (1982), baseando-se em propostas feitas anteriormente por Pike & Pike (1947) e Fudge (1969).

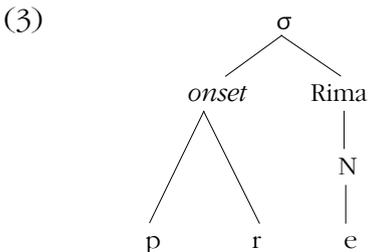
(1)



Na estrutura silábica do português, somente o núcleo é obrigatório, pois a língua pode apresentar sílabas com *onset* e/ou coda zero. O *onset* da primeira sílaba da palavra não pode ser constituído pelos segmentos /t/, /ɲ/, /ʎ/. A estrutura do *onset* pode possuir até duas consoantes, sendo a primeira uma plosiva ou fricativa labial, e a segunda, sempre uma líquida, como os exemplos em (2a). O núcleo silábico no português só pode ser formado por vogais. A estrutura da coda somente aceita /N/, /S/, /l/, /r/, como se pode verificar em (2b), ou também pode apresentar duas consoantes, sendo que a segunda deve ser sempre um /S/, conforme se pode ver em (2c).

- (2)
- |     |                                        |                 |
|-----|----------------------------------------|-----------------|
| (a) | prato ['pratʊ]                         | fraco ['fraku]  |
|     | bloco ['blɔku]                         | flor ['flɔr]    |
| (b) | campo ['kampu]                         | neste ['nestʃi] |
|     | sol ['sɔɫ] ~ ['sɔw] <sup>1</sup>       | marca ['markə]  |
| (c) | perspicaz [perspi'kas]                 |                 |
|     | solstício [soɫs'tisju] ~ [sows'tʃisju] |                 |

A inclusão de mais de um segmento na estrutura do *onset*, o que torna o *onset* ramificado, conforme podemos ver em (3), confere um grau de maior complexidade a essa estrutura. O inventário silábico do português permite um número limitado de encontros consonantais tanto em *onset* quanto em coda: (C) (C) V (C) (C).



Assim como na estrutura silábica do português, no inglês, também, somente o núcleo é obrigatório, podendo apresentar sílabas com *onset* e/ou coda zero.

Em inglês, a estrutura CCV apresenta dois tipos: a sílaba começa com a fricativa /s/, seguida de consoantes plosiva surdas, fricativa labial surda, nasais anteriores, líquida lateral e glides, conforme se mostra nos exemplos em (4a); ou a sílaba pode iniciar por consoantes plosivas, fricativas, africadas, nasais anteriores e líquidas seguidas de glides e líquidas, como aparece nos exemplos em (4b). A coda pode apresentar uma consoante soante ou uma obstruente, conforme aparece exemplificado em (4c). A estrutura silábica do inglês, diferentemente do português, permite vários encontros consonantais: (C) (C) (C) V (C) (C) (C) (C).

|     |     |                  |                     |
|-----|-----|------------------|---------------------|
| (4) | (a) | speak ['spi:k]   | stool ['stu:l]      |
|     |     | sleep ['sli:p]   | sfumato [sfu'matou] |
|     |     | smart ['sma:rt]  | snow ['snou]        |
|     | (b) | plum ['plʌm]     | bread [ˈbrɛd]       |
|     |     | flower [ˈflaʊər] | friend [ˈfrend]     |
|     | (c) | car [ˈca:r]      | pen [ˈpen]          |
|     |     | gum [ˈɡʌm]       | stop [ˈstɑ:p]       |
|     |     | paint [ˈpeɪnt]   | clock [ˈklɑ:k]      |

No português, diferentemente do inglês, não há ocorrência de *onset* complexo que comece com /s/ no início da palavra. Portanto, determinou-se, como objeto desta pesquisa, o estudo de encontros consonantais iniciados por /s/ em início de palavra, produzidos por uma criança que está adquirindo o português como LM e o inglês como LE. Uma vez que se sabe que mesmo em palavras consideradas empréstimos, como por exemplo *stress*, os falantes nativos do português brasileiro tendem a produzi-la [is'trɛs], com um [i] epentético. Esse *output* é obtido porque o falante nativo do português brasileiro, ao se deparar com uma estrutura não existente em sua língua, lança mão de uma estratégia, nesse caso estratégia de epêntese, para tornar essa estrutura compatível com as estruturas de sua LM.

Os estudos sobre aquisição de língua estrangeira têm apontado a influência da língua materna nesse processo. Em uma pesquisa com adolescentes e adultos, Fernandes (2001) e Silveira (2002) mostram que, no processo de aquisição do inglês como língua estrangeira por

falantes nativos de português, a LE sofre interferência da estrutura silábica da língua materna, já que essas duas línguas apresentam algumas estruturas silábicas diferentes. Em se tratando de crianças, porém, essa interferência da LM na aquisição da LE parece ser menor. De acordo com Fletcher e MacWhinney (1997, p.186), a criança que é exposta ainda cedo a uma segunda língua desenvolve mais precocemente uma consciência sobre a linguagem, o que pode diminuir essa influência da LM sobre uma LE em aquisição.

A consciência metalingüística é a habilidade de pensar sobre a linguagem como um objeto e desenvolve-se gradualmente durante a infância, só estando totalmente estabelecida aos oito ou nove anos de idade. As crianças pequenas, como os falantes em geral, normalmente tendem a ver a linguagem como um meio de comunicação, com ênfase principal no conteúdo e no uso e não na forma de um enunciado (FLETCHER e MACWHINNEY, 1997, p.288). Segundo Morais (1989), a consciência fonológica é uma capacidade metalingüística que se refere à representação consciente das propriedades fonológicas e das unidades constituintes da fala, ou seja, é a consciência dos sons que compõem as palavras que ouvimos e falamos (CARDOSO-MARTINS, 1991). A capacidade metafonológica pressupõe o conhecimento não do que significa uma palavra, mas de que ela é um todo formado por unidades fonológicas menores, como por exemplo fonemas e sílabas. A consciência fonológica requer que a criança ignore o significado e preste atenção à estrutura da palavra (GOUGH e LARSON, 1996, p.15).

Esta pesquisa tem como objetivos, através de um estudo do processo de aquisição de inglês como LE por uma criança em etapa de desenvolvimento fonológico, verificar o processo de aquisição de encontros consonantais do inglês começados por /s/ no início de palavra e, também, verificar a existência de transferência de estruturas silábicas da LM para a LE.

Tais objetivos advêm da hipótese sobre a existência de transferência de estruturas silábicas da LM para a LE, mesmo em etapa precoce de aquisição da fonologia, e também da hipótese de que, assim como o adulto, a criança utiliza a estratégia de epêntese para reestruturar

a sílaba ao padrão universal CV, ou a estruturas licenciadas em sua língua materna.

Para a realização da pesquisa, determinou-se um *corpus* constituído de dados de uma criança em desenvolvimento fonológico normal, no período em que se encontrava com idade de 3:4 (anos/meses). A criança freqüentava escola maternal no município de Pelotas, Rio Grande do Sul, e pertence ao nível sócio-econômico-cultural de classe média. Está adquirindo o português como língua materna e recebe da mãe, também nativa de Português, o *input* em inglês desde a idade de 2:1.

Foi analisado um *corpus* constituído por quarenta e cinco palavras em inglês que apresentam *onset* complexo iniciado por /s/ na borda esquerda da palavra. Os encontros consonantais iniciados em /s/ podem ser do tipo duplo, quando, além do /s/, apresentar mais uma consoante, ou do tipo triplo, quando for constituído por /s/ e mais duas consoantes. Os dados para esta pesquisa são constituídos por esses dois tipos de encontros consonantais. Para a coleta de dados, aplicou-se a técnica de tradução eliciada, através da qual, depois de ouvir determinada palavra em português, a criança é estimulada a produzi-la em inglês, com o objetivo de se obter uma amostra lingüística sem imitação. Algumas vezes foi necessário o uso da técnica de imitação retardada, isto é, a produção de uma palavra em uma frase, sucedida de uma série de outras palavras, e, depois, é feita a solicitação de sua realização pela criança. Para que se possa obter uma amostragem significativa, é necessário que a criança já possua um vasto vocabulário na língua estrangeira.

Todos os dados coletados foram transcritos foneticamente. Após a transcrição fonética das fitas cassetes, os dados foram analisados para obter-se a forma como essas estruturas silábicas foram produzidas pela criança.

A menina, cuja fala originou o *corpus* que embasa este estudo, apresentou em sua produção, com exceção de /spl/, todos os tipos de encontros consonantais iniciados por /s/ licenciados na língua inglesa, como se verifica nos exemplos em (5).

|     |      |        |            |       |         |           |
|-----|------|--------|------------|-------|---------|-----------|
| (5) | /sl/ | sleep  | ['sli:p]   | /skr/ | scream  | ['skrim]  |
|     | /sm/ | small  | ['ismɔ:l]  | /str/ | straw   | ['strɔ:]  |
|     | /sn/ | snake  | ['snejk]   | /skw/ | squeeze | ['skwiz]  |
|     | /sp/ | spider | ['spajdər] | /skr/ | scream  | ['skrim]  |
|     | /st/ | stool  | ['stul]    | /spr/ | spring  | ['sprɪŋg] |

Sabe-se que estruturas mais marcadas possuem um grau de dificuldade maior que estruturas menos marcadas, portando o *onset* constituído por dois elementos é mais facilmente adquirido por falantes, que o constituído por três elementos.

Pesquisas têm indicado que as duas estratégias mais utilizadas para simplificação de estruturas silábicas, presentes na LE que não seguem o padrão da sílaba da LM, são epêntese e apagamento. Para falantes nativos de inglês a estratégia de apagamento é a preferida, enquanto para os falantes nativos de português aprendendo inglês, a epêntese é mais utilizada.

A epêntese é a estratégia comprovadamente mais usada por adultos nesse tipo de *onset* complexo (FERNANDES, 2001 e SILVEIRA, 2002); e também foi a estratégia usada pela menina informante da presente pesquisa.

Os dados da presente pesquisa confirmam essa preferência de falantes nativos de português fazerem uso de epêntese ao utilizar inglês, pois a criança, cujos dados serviram de base para este estudo, utilizou essa estratégia para simplificar 15,55% dos encontros consonantais iniciados por /s/ em inglês, conforme se verifica em (6).

(6)

|                     | 100%<br>de epêntese<br>(4/4) | 40%<br>de epêntese<br>(2/5) | 25%<br>de epêntese<br>(1/4) | 84,44%<br>de realizações<br>apropriadas<br>(38/45) |
|---------------------|------------------------------|-----------------------------|-----------------------------|----------------------------------------------------|
| <i>onset</i> duplo  | [sm]                         | [sn]                        |                             | [sk, sl, sn, sp, st, str]                          |
| <i>onset</i> triplo |                              |                             | [str]                       | [skr, spr, skw]                                    |

O que os dados em (6) estão mostrando, é que, embora a criança faça uso de epêntese, essa estratégia está preferencialmente

sendo empregada quando o encontro consonantal é constituído de /s/ seguido de uma nasal, conforme aparece exemplificado em (7). Fica evidente também que a estrutura '/s/+nasal' está em processo de aquisição pela criança, pois, embora o encontro /sm/ tenha sido produzido com 100% de epêntese, o *onset* formado por /sn/ obteve 60% de suas realizações apropriadas.

|     |         |           |       |           |
|-----|---------|-----------|-------|-----------|
| (7) | snack   | ['isnek]  | small | ['ismal]  |
|     | snake   | ['isnejk] | smart | ['ismart] |
|     | sneaker | ['snikər] | smell | ['ismel]  |
|     | sneeze  | ['sniz]   | smile | ['ismajl] |
|     | snow    | ['snow]   |       |           |

Pelos dados em (6) também pode-se observar que a menina apresenta maiores problemas no emprego de encontros consonantais iniciados por /s/ constituídos de dois elementos (83,33% de realizações apropriadas) do que em *onset* complexo constituído de três elementos (88,88% de realizações apropriadas). Diante desse resultado, tem de questionar-se a causa desse fato, que pareceria contrariar o encaminhamento de a aquisição dar-se do menos marcado para o mais marcado.

A resposta pode estar no grau de 'sonoridade' dos segmentos constitutivos dos *onsets* complexos do inglês. A escala de sonoridade possui um importante papel na estrutura silábica, pois relaciona a sonoridade de um segmento com a posição que ele ocupa no interior da sílaba. O elemento mais sonoro sempre ocupa o núcleo da sílaba, e os menos sonoros ocupam o *onset* e a coda. Quando o *onset* ou a coda são complexos, a sonoridade de seus segmentos é crescente em direção ao núcleo, de acordo com a escala de sonoridade, conforme apresentada em (8).

|     |       |   |       |   |         |   |       |   |           |   |          |   |         |
|-----|-------|---|-------|---|---------|---|-------|---|-----------|---|----------|---|---------|
| (8) | vogal | > | glide | > | líquida | > | nasal | > | fricativa | > | africada | > | plosiva |
|     | 6     |   | 5     |   | 4       |   | 3     |   | 2         |   | 1        |   | 0       |

A sonoridade do *onset* é sempre crescente, se se refere ao pico silábico, e a da coda, decrescente. O princípio do ciclo de sonoridade,

segundo Clements (1990), diz que a curva de sonoridade aumenta maximamente do *onset* para o núcleo e cai minimamente do núcleo para a coda. Na constituição da sílaba, também há o princípio da distância de sonoridade, que trata da relação de sonoridade entre os segmentos que integram o constituinte complexo. Segundo esse princípio, para o português e espanhol, os elementos do encontro consonantal têm que ter uma distância mínima de dois pontos (HARRIS, 1983).

Empregando-se o princípio da distância de sonoridade aos encontros consonantais que fazem parte do *corpus* deste trabalho, obtêm-se os resultados mostrados em (9).

(9)

(a) Encontros consonantais duplos

|                   | /sk/² | /sl/ | /sm/ | /sn/ | /sp/ | /st/ |
|-------------------|-------|------|------|------|------|------|
| Valor na escala   | 2 0   | 2 4  | 2 3  | 2 3  | 2 0  | 2 0  |
| Grau de distância | 2     | 2    | 1    | 1    | 2    | 2    |

(b) Encontros consonantais triplos

|                   | /skw/ | /skr/ | /spr/ | /str/ |
|-------------------|-------|-------|-------|-------|
| Valor na escala   | 2 0 5 | 2 0 4 | 2 0 4 | 2 0 4 |
| Grau de distância | 2 5   | 2 4   | 2 4   | 2 4   |

Conforme podemos observar em (9a), os únicos encontros consonantais que não obedecem ao princípio da distância de Sonoridade de dois pontos entre seus constituintes – que é o grau exigido no sistema do português –, são os formados por ‘/s/+nasal’. Esse fato pode explicar por que esse tipo de *onset* é o único grupo consonantal para cuja produção a criança apresenta dificuldade.

Em todos os casos de emprego de epêntese, a informante utilizou a vogal [i] como elemento epentético. Esse fato não causa surpresa, uma vez que [i] é a vogal epentética por excelência do português. A epêntese em palavras iniciadas por /s/ resulta em sílabas do tipo VC, que, embora não constitua um universal lingüístico, é uma estrutura presente na língua materna da falante e também é uma sílaba de

aquisição precedente, por crianças brasileiras, em se comparando com a estrutura CCV (MATZENAUER, 1990; LAMPRECHT, 1990).

No caso dos encontros formados por três consoantes, apesar de serem estruturas mais marcadas que os encontros duplos, e portanto mais difíceis de serem produzidas, já foram adquiridos pela criança. Conforme se pode ver em (6), somente o *onset* do tipo /str/ apresenta caso de epêntese, conforme se mostra nos exemplos em (10a). Essa estratégia somente foi utilizada quando após o encontro /str/ há a presença da vogal /i/, como aparece nos exemplos em (10b), ou seja, a mesma vogal que é utilizada como segmento epentético no português; esse núcleo silábico parece estar favorecendo a epêntese, em caso de *onset* constituído por três elementos.

(10)

|     |       |            |              |
|-----|-------|------------|--------------|
| (a) | /skr/ | scratch    | [ˈskretʃ]    |
|     | /spr/ | spring     | [ˈsprɪŋg]    |
|     | /skw/ | square     | [ˈskwɛr]     |
|     | /str/ | straw      | [ˈstrɔ]      |
|     | /str/ | street     | [ˈɪstrɪt]    |
| (b) | /str/ | straw      | [ˈstrɔ]      |
|     | /str/ | strawberry | [ˈstrɔ:berɪ] |
|     | /str/ | street     | [ˈɪstrɪt]    |
|     | /str/ | strong     | [ˈstrɔŋg]    |

Verificando-se, portanto, que o princípio de distância de Sonoridade está atuando no processo de aquisição de encontros consonantais iniciados por /s/ na informante deste trabalho, cabe-se ainda questionar por que essa aquisição está ocorrendo inicialmente com /sn/ que já apresenta 60% de realizações adequadas, enquanto a seqüência /sm/ mostra 100% de emprego de epêntese.

Como, nesse caso, a distância de sonoridade entre os dois segmentos do encontro é a mesma, é preciso analisar a estrutura interna de cada elemento, ou seja, os traços distintivos constituintes de cada segmento. Pode-se ver, então, que, na seqüência /sn/, os dois segmentos compartilham o ponto de articulação [coronal], enquanto na seqüência /sm/, o primeiro elemento tem o ponto [coronal] e o segundo, o ponto [labial].

Pelos dados em (6), a menina está adquirindo, em uma primeira etapa, os encontros com segmentos que compartilham o ponto de articulação, que é o caso de /sn/. Em uma etapa subsequente, deverão emergir em seu sistema fonológico os encontros com segmentos que têm pontos de articulação diferenciados, que é o caso de /sm/.

Esse resultado está apontando para a ação de outro princípio, durante o processo de aquisição de encontros consonantais: o princípio de co-ocorrência de traços. Esse princípio trata de co-ocorrência de traços de segmentos adjacentes (esse fenômeno é conhecido na literatura como OCP – Obligatory Contour Principle).

Já foi amplamente comprovado pela literatura da área que, quanto antes uma LE for adquirida, melhor será o desempenho do aprendiz no que se refere à pronúncia. Quanto mais desenvolvido estiver o sistema fonológico da LM ao iniciar o aprendizado de uma LE, menor vai ser sua habilidade em perceber e adquirir os sons da LE (FLEGE, 1999). Pode-se concluir, então, que a criança parece perceber mais rapidamente sons que não pertencem a sua língua materna.

Pela pesquisa aqui realizada, vê-se que os dados parecem vir ao encontro da afirmação de Flege (1999) no sentido de que crianças expostas a uma segunda língua, ainda no período em que estão adquirindo a LM, apresentam produção adequada em período de tempo menor que o adulto; isto quer dizer que, quanto menor for a criança, há menos interferência da LM.

Comparando-se os dados da menina estudada neste trabalho com os dados de Silveira (2002) sobre a percepção e produção de encontros consonantais do inglês iniciados por /s/ por aprendizes brasileiros adultos, pode-se verificar que no período de 1 ano e três meses, entre 2:1 e 3:4, a criança produz 84,44% dos encontros consonantais iniciados por /s/ de forma adequada. Pelos dados de Silveira (2002), alunos adultos brasileiros, com um ano e sete meses e 4 anos de estudo de inglês como LE, ainda apresentam dificuldade na produção de *onsets* iniciados por /s/, mostrando a utilização de uma vogal epentética na produção de todos os encontros iniciados por /s/.

Conclui-se, portanto, que a criança ainda em fase de aquisição da fonologia parece mostrar maior facilidade e rapidez no domínio de estruturas marcadas – aqui se trata especificamente de *onsets* complexos constituídos por duas ou três consoantes iniciados por /s/ em início de palavra – da LE do que adultos.

A autonomia, no sentido de estabelecimento de estratégias individuais, existe na aquisição de LE, mas é regulada também pelo sistema que a criança já tem internalizado, que é o da LM. Nesse sentido, a LM, ao influenciar o processo de aquisição de uma nova língua, mostra-se como limitadora da ‘autonomia’ – a criança estudada não faz apagamentos (que falantes nativos de inglês fazem), faz epênteses, característica da LM.

## NOTAS

<sup>1</sup> O /l/ em coda é semivocalizado em algumas variantes lingüísticas do PB.

<sup>2</sup> Nas seqüências com encontros em que /s/ antecede uma plosiva, há a distância de dois graus de sonoridade, mas também há uma reversão de sonoridade, violando o princípio de seqüência de sonoridade (CLEMENTS, 1990, p.285) mas, segundo Clements (1990, p. 311-12), o comportamento especial das coronais anteriores não deve ter relação com a escala de sonoridade.

## REFERÊNCIAS BIBLIOGRÁFICAS

CARDOSO-MARTINS, C. A sensibilidade fonológica e a aprendizagem inicial da leitura e da escrita. *Cadernos de pesquisa*, v. 76, p. 41-49, fev. 1991.

CLEMENTS, G. N. The role of the sonority cycle in core syllabification. In: KIGSTON, J.; BECKMAN, M. (Org.). *Papers in laboratory phonology 1*. Cambridge: CUP, p. 283-333, 1990.

FERNANDES P. R. C. A epêntese nas formas oral e escrita na interfonologia Português/Inglês. In: MATZENAUER, C. L. (Org.). *Aquisição de língua materna e de língua estrangeira*. Aspectos fonético-fonológicos. Pelotas: EDUCAT, 2001. p. 235-259.

FLEGE, J. Age of learning and second language speech. In: BIRDSONG, D. (Ed.). *Second language acquisition and the critical period hypothesis*. Mahwah: Lawrence Erlbaum Associates, 1999.

- FLETCHER, P.; MACWHINNEY, B. *Compêndio da linguagem da criança*. Porto Alegre: Artes Médicas, 1997.
- FUDGE, E. Syllables. *Journal of Linguistics*, Cambridge, UK, n. 5, p. 254-287, 1969.
- GOUGH, P.; LARSON, K. A estrutura da consciência fonológica. In: CARDOSO-MARTINS, C. *Consciência fonológica e alfabetização*. Petrópolis: Vozes, 1996.
- HARRIS, J. *Syllable structure and stress in Spanish*. A non linear analysis. Cambridge: MIT Press, 1983.
- INGRAM, D. *First language acquisition: method, description and explanation*. USA: Cambridge, 1989.
- LAMPRECHT, R. R. *Perfil de aquisição normal de fonologia do português – Descrição longitudinal de 12 crianças: 2:9 a 5:5*. 1990. Tese (Doutorado em Letras) – Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.
- MATZENAUER, C. L. *Aquisição da fonologia do português: estabelecimento de padrões com base em traços distintivos*. 1990. Tese (Doutorado em Letras) – Instituto de Letras e Artes, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.
- MORAIS, J. Phonological awareness: a bridge between language and literacy. In: SAWYER, D. J.; FOX, B. *Phonological awareness in reading: the evolution of current perspective*. Berlin: Springer, 1989.
- PIKE, K.; PIKE, E. Immediate constituents of Mazateco syllables. *International Journal of Applied Linguistics*, n. 13, p. 78-91, 1947.
- SELKIRK, E. The syllable. HULST, H. V. D., SMITH. *The structure of phonological representations* (part. II). Foris, Dordrecht, p. 337-383, 1982.
- SILVEIRA, R. Perception and production of English initial /s/ clusters by Brazilian learners. *Revista Brasileira de Linguística Aplicada*, n. 1, p. 95-119, 2002.

# **Interpretação fonológica e realização fonética durante o aprendizado de uma LE**

Nina A. Lyubimova<sup>1</sup>

**Professora Visitante CNPq/Instituto de Letras, UFRGS**

Tradução de Tanira Castro<sup>2</sup>

## **Abstract**

In the process of learning a foreign language, the individual must master a new phonetic discourse formation and its respective reception in that language, according to the “rules” of the intermediary system of paradigmatic units which are not similar to phonemes either of the first or the second system. These special units have a mixed character that differentiates them from a free variation. They can be called diaphonemes, if we understand that these units, which have a double character, become manifest in relation to the intermediary linguistic system and show themselves to be paradigmatic units. Nevertheless, in relation to the system of language that is studied, they show themselves to be syntagmatic units, and their variations are the realizations of the phonemes of the idiom in question.

O discurso sonoro representa o elo principal da comunicação acústica do indivíduo e, por isso, é reconhecido como um dos principais componentes da atividade comunicativa de conhecimento. Nesse sentido, os sons do discurso representam o principal objeto do ensino e da aprendizagem da LE. Para este artigo, temos em mente que não apenas o som completo representado por um determinado “conjunto” de movimentos dos órgãos articulatórios da fala, mas também o fato de que ele é o portador de uma determinada informação acústica, que pode ser representada pelo tom, timbre, intensidade, duração –, mas também o significado absoluto desse som, assim como as unidades segmentais sejam secundários, mas as suas variações de seus significados relativos representam precisamente a essência da fonética, como, por exemplo, a tonicidade e a entonação. Naturalmente, as informações gramaticais e semânticas também não podem ser atualizadas se esquecermos os sons do discurso.

O domínio do discurso em uma LE necessariamente pressupõe a formação de novas informações automatizadas que garantam a percepção e a realização das unidades de diferentes níveis dos componentes fonológicos de uma determinada LE, de acordo com o seu sistema e normas. A formação de novas operações ocorre sob as condições da ação da interferência interlingual, isto é, em condições de uma mútua interação de dois, e, às vezes mais, sistemas fonéticos (num sentido amplo) na consciência do indivíduo. Em consequência à ação da interferência, de uma forma específica, manifesta-se uma contradição entre os critérios universais e específicos, entre as semelhanças e diferenças existentes no primeiro ou no segundo sistemas lingüísticos e, por isso, o resultado dessa ação pode ser tanto

positivo quanto negativo. De acordo com a interferência interlingual, possuímos também a interferência das operações auditivas motoras discursivas, formadas durante o processo de aprendizagem da LE e operações formadas durante o domínio de uma nova linguagem. A ação de mútua interação desta ou daquela interferência significativamente age durante o aprendizado do aspecto fonético do discurso da língua que se estuda.

Como sabemos, o indivíduo percebe o discurso na língua que estuda através do prisma de suas associações auditivas, que se baseiam nas línguas antes estudadas. Entretanto, observamos que a resposta à questão de qual das línguas que antes foi aprendida apresenta maior grau de interferência manifesta na aprendizagem da nova língua, permanece até agora uma questão que deixa muitas dúvidas para discussão.

No discurso em sua língua materna, o indivíduo é capaz de identificar os diferentes sons físicos como uma única unidade lingüística, porque as suas operações auditivas já estão automatizadas e obedecem às relações fonológicas de tal língua. Por outro lado, como sabemos, o indivíduo é capaz de diferenciar um maior número de sons do que a qualidade dos fonemas em sua língua materna (BONDARKO, 1981) e como eles se manifestam, o que representa a razão da identificação incorreta das características fonêmicas e dos fonemas em sua totalidade, durante a percepção dos sons da outra língua. Durante o domínio incompleto (pequeno) da língua estudada, o indivíduo percebe as diferenças fonêmicas, baseando-se em associações incorretas, em virtude das semelhanças físicas de seus correlatos “próprios” e “alheios” das características fonêmicas. Isso provoca, conseqüentemente, que em um mesmo nível fonético, também aquele mesmo critério diferencial possui diferentes características objetivas (ZINDER, 1979).

Se considerarmos não a etapa inicial de domínio de uma língua, mas o bilingüismo subordinado em diferentes situações de comunicação, então deveremos reconhecer que a ação da interferência fonética, durante a percepção do discurso na LE, normalmente

permanece fechado para os interlocutores. O fato de que, mesmo num determinado nível de domínio da LE, o indivíduo cometa erros fonéticos, não impede a correta percepção do conteúdo significativo do discurso do interlocutor nativo de uma determinada língua, ainda não significa que, decodificando o sinal sonoro ele utiliza o sistema de critérios lingüísticos de tal LE. A ação negativa na interferência fonética em um determinado grau conhecido neutraliza o comportamento e a orientação significativa do discurso do ouvinte e diferentes tipos de informações supérfluas nele contidas. Além disso, durante o bilingüismo subordinado, o indivíduo em algum determinado grau já tem formado o seu sentido de adivinhação lingüística na outra língua, isto é, com maior ou menor sucesso, atua um mecanismo de prognóstico possível, permitindo um reconhecimento correto por parte do ouvinte. Ao contrário, palavras isoladas que contenham dificuldades fonéticas para portadores de outras LE, como de regra, pode perceber e identificar corretamente. Para identificar a ação da interferência, às vezes, é bastante difícil porque o indivíduo pode não prestar atenção nos erros do seu próprio discurso, mas fixar erros análogos no discurso de outras pessoas, que também não dominam completamente o aspecto sonoro do discurso da “outra” língua. Isso testemunha o fato de que o mecanismo auditivo de relação inversa ainda não está completo na LE, apesar de determinados padrões já existirem, o que permite a percepção dos erros no discurso do interlocutor. É possível também a seguinte situação: quando o indivíduo conhece e identifica os seus erros de pronúncia, consegue ouvi-los, mas não se encontra em condições de se autocorrigir porque a formação e a automatização das operações discursivas motoras na LE significativamente atrasaram-se em relação à formação e à automatização das operações auditivas. Observamos que situações semelhantes também se manifestam na LM, não somente na LE (BERNSHTEIN, 1930; JAKOBSON, 1968).

A percepção categorial em uma nova língua chamada de audição fonológica, forma-se simultaneamente com o programa motor que responde não somente pela articulação de um som concreto. Ela

também responde por quais órgãos discursivos participam na formação das diferentes características da LE e com qual intensidade, quais as combinações que funcionam no trabalho articulatório dos órgãos da fala e que são possíveis em tal LE (ZINDER, 1979, p. 81). Os psicólogos, neste caso, falam sobre os mecanismos de fonação ou os mecanismos de formação externa da comunicação. Durante o ensino/aprendizagem de uma LE, tais mecanismos funcionam de uma forma diferente da maneira como eles funcionam de acordo com os programas da LM, isto é, aqui tem lugar uma seqüência de comandos silábicos motores dos movimentos articulatórios (ZIMNIAIA, 1989).

O sistema de correspondências sonoras que surge durante a interferência, não permite ao indivíduo reproduzir corretamente as formas sonoras de uma determinada palavra durante a produção do discurso e tão pouco permite a percepção adequada da palavra, baseando-se apenas em suas características objetivas, de acordo com o sistema da LE em pauta. A explicação disto deve-se à passagem de transição dos sistemas fonéticos interagentes que, sem levar em consideração a sua labialização, a sua má estruturação, falando metaforicamente, forma uma espécie de composição “híbrida”, que se identifica como inicial e assim, a partir do segundo sistema sonoro, possui as suas características sintagmáticas, paradigmáticas e diferenciais (LYUBIMOVA, 1991).

O discurso incorreto dos estrangeiros durante o processo de aprendizagem de uma determinada LE, pelo menos no que diz respeito ao seu aspecto fonético incorreto, é mais uma manifestação dessa variação dentro dos limites dos sistemas da LE, mas que ocorre em virtude da ação da interferência durante as condições da formação do bilingüismo. A variação fonética do discurso de estrangeiros pode extrapolar os limites das propriedades do sistema lingüístico de uma determinada LE (o subsistema), formando, dessa forma, o seu próprio sistema (outro subsistema). Mas o mesmo pode coincidir com algumas características de algum dialeto ou algum tipo de falares, o que abre caminho para um outro sistema, que pode não ter nada a ver com o primeiro sistema. Naturalmente que tal sistema pode variar, não

somente no discurso de aprendizes de diferentes idiomas, mas também no discurso de portadores de um mesmo idioma. Dessa mesma maneira, pode variar também o número de indivíduos que utilizam esta ou aquela variante.

Nesse caso, podemos constatar que existe algo em comum, por exemplo, com aquilo que está presente a todos os russos que não dominam a língua portuguesa. Aqui, os sistemas lingüísticos individuais, obrigatoriamente, interagem, isto é, possuem alguma determinada parte em comum que nos permite falar a respeito do sotaque russo no discurso na língua Portuguesa. Mas os sistemas intermediários e os interagentes são os primeiros sistemas representados por diferentes línguas até mesmo se entre esses sistemas não existe nenhuma semelhança genética ou tipológica. Desta forma, sabemos bem que os falantes de diferentes idiomas podem cometer os mesmos erros fonéticos no discurso na língua russa. Um exemplo disso pode servir a violação das oposições das consoantes surdas e sonoras, duras e brandas, cometidas por indivíduos falantes de diferentes línguas.

As pesquisas necessárias para o trabalho da articulação e a realização desta ou daquela característica diferencial, em cada caso concreto, sem dúvida nenhuma, possui “as suas características”. Isso porque refletem não somente o “conflito” sistêmico, mas também o “conflito” do trabalho mecânico de fonação, formado anteriormente, em relação ao que está se formando, em resultado das tentativas por parte do aprendiz, em produzir o discurso normativo fonético na LE em questão. Desta forma, temos aqui a violação das oposições fonêmicas e as violações das variações alofônicas. Assim, devemos observar que em uma determinada etapa do domínio de alguma LE, a composição fonêmica dos sistemas interagentes, em um grau significativo, coincide, porque supõe-se que, em condições de um bilingüismo subordinativo, o sistema intermediário baseia-se, antes de mais nada, nos componentes universais da linguagem, entre eles os sistemas fonéticos e demais sistemas. As diferenças entre esses sistemas, de acordo com as suas semelhanças tipológicas genéticas, devem ser expressas principalmente através do caráter variativo dos seus

alofones e suas correspondências com os fonemas da LE que se estuda.

A descrição fonética da formação do discurso em uma LE, até hoje, limita-se à descrição de um determinado número de realizações fonêmicas. A identificação das funções fonológicas e dos sons, com a qual o indivíduo trabalha durante o domínio incompleto de uma LE, apresenta determinadas dificuldades, mencionadas inúmeras vezes na literatura que aborda tal assunto. Realmente, por um lado, nós conhecemos as unidades que os sons representam – que são os fonemas de uma determinada língua. Por outro lado, reconhecer isso ainda não nós é possível, pois a realização concreta de um determinado fonema contém um determinado som, que corresponde ou a um fonema do primeiro sistema, ou a um fonema de um outro sistema lingüístico, ou, finalmente, esse fonema não é característico nem do primeiro e nem do segundo sistema. Além disso, não podemos deixar de reconhecer que, normalmente, todas as análises realizadas nesse sentido evidenciam uma assimetria na realização das características fonêmicas durante a percepção ou a produção do discurso. Tais características demonstram, como já dissemos acima, o fato de que o indivíduo é capaz de distinguir uma grande quantidade de fonemas, apesar de não se encontrar em condições ativas de utilizá-los no discurso.

Neste caso, a questão pode relacionar-se a fonemas com múltiplos alofones, cuja unidade deve-se à identificação das palavras, nas quais eles podem apresentar-se em qualquer uma das variantes facultativas (ZINDER, 1979, p. 73). Tal tipo de unidade podemos chamar de *diafonema* (LYUBIMOVA, 2001), compreendendo que o conceito dessa unidade possui um caráter duplo, que se manifesta em relação ao sistema intermediário, onde ela se apresenta como uma unidade paradigmática. Entretanto, em relação ao sistema de uma determinada língua, ela pode apresentar-se com um caráter sintagmático, pois as suas variantes são a realização dos fonemas dessa língua. Lembramos que já falamos a respeito do caráter duplo sistêmico dessas unidades, durante o domínio incompleto do aspecto

fonético do discurso na LE (HAUGEN, 1956; ROZENTSVEIG, 1972). Entretanto, este caráter duplo sistêmico normalmente é analisado como uma característica de ambos os sistemas lingüísticos em contato. Em relação aos aspectos paradigmáticos e sintagmáticos desta questão, salientamos que os mesmos ainda não foram analisados suficientemente, pois o termo “diafonema” até hoje não foi utilizado. Diante do exposto, o termo “diafone”, de acordo com a teoria de Jones (1950) e de acordo com a teoria de E. Haugen (1954), não são idênticos ao conceito de “diafonema”. Os componentes das variantes de um *diafonema* podem coincidir, em primeiro lugar, com os alofones de fonemas semelhantes do primeiro sistema; em segundo lugar, com os alofones de diferentes fonemas do segundo sistema. Mas devemos levar em consideração que são possíveis composições mais complexas: alofones de diferentes fonemas do segundo sistema e alofones de fonemas semelhantes do primeiro sistema; ou ainda uma composição múltipla, que inclui alofones desconhecidos, tanto no primeiro sistema quanto no segundo. Evidentemente, tais tipos de unidades são possíveis e pode haver uma coincidência total com os alofones de fonemas do segundo sistema.

A veracidade de tal conceituação pode ser demonstrada através dos resultados de um amplo círculo de pesquisas baseadas em materiais de diferentes línguas em contato, que consideram não somente o aspecto segmental, mas também o nível supra-segmental dos componentes fonológicos dos idiomas em questão.

## NOTAS

<sup>1</sup> Nina A. Lyubimova – Doutora em Ciências Filológicas, áreas Lingüística Geral e Teoria e Metodologia do Ensino de Língua Estrangeira. Professora Titular da Faculdade de Filologia, Universidade Federal de São Petersburgo, São Petersburgo – Rússia.

<sup>2</sup> Tanira Castro – Professora Adjunto do Setor de Russo, Departamento de Línguas Modernas – LET2, Instituto de Letras – UFRGS.

## REFERÊNCIAS BIBLIOGRÁFICAS

BERNSHTEIN, S.I.<sup>1</sup> *Voprosy obutchenia proiznosbeniu (primenitelno k prepodavaniiu russkogo iazyka inostrantsam) // Voprosy fonetiki i obutchenie proiznosheniu*. M.: MGU, 1975. *Questões relativas ao ensino-aprendizagem da pronúncia (aplicadas ao ensino da língua russa para estrangeiros)*. // *Questões relativas à fonética e ao ensino-aprendizagem da pronúncia*. Moscou: ed. da UFMoscou. 1975.

BONDARKO, L. V. *Foneticheskoe opisanie iazyka i fonologicheskoe opisanie retchi*. L.: LGU, 1981. *Descrição fonética da linguagem e descrição fonológica do discurso*. Leningrado: Ed. LGU. 1981.

HAUGEN, E. *Bilingualism in the Americas. A Bibliography and Guide to Research*. Publications of the American Dialect Society University of Alabama 26, 1956.

HAUGEN, E. *Problems of Bilingualism Description*. Georgetown University Monograph Series on Language and Linguistics 7, September 1954, p. 9-19.

JONES, D. *The Phoneme: Its Nature and Use*. Cambridge, Heffner1950 (2<sup>nd</sup> ed. 1962; 3<sup>rd</sup> ed. 1967).

JAKOBSON, R.O. *Child Language, Aphasia and Phonological Universalis*. The Hague; Paris: Mouton, 1968.

LYUBIMOVA, N.A. *Foneticheskaja interpfერentsya i obsnenie na nerodnom iazyke (experimentalno-foneticheskoe issledovanie na materiale finsko-russkogo dviiazytchia)*. SPb.: UFSPb, 1991. Ed. *A interferência fonética e a comunicação na LE (pesquisa de fonética experimental baseada em material de bilingüismo russo-finlandês)*. Resumo de tese de doutorado. São-Petersburgo: Ed. UFSPb, 1991.

LYUBIMOVA, N.A. *On phonetic Language Interference and the Intermediate System*. // 100 Years of Experimental Phonetics in Russia. S.- Petersburg: Ed. UFSPb, 2001. P. 127-130.

ROZENTSVEIG, V. I. *Iazykovyie kontakty. Lingvisticheskaja problematika*. L.: Nauka, 1972. *Contatos lingüísticos. Questões de lingüística*. Leningrado: Ed. Nauka. 1972.

WEINREICH, U. *Languages in Contact: findings and problems*. Linguistic Circle of New York, New-York, 1953.

ZIMNIAIA, I.A. *Psikhologia obutchenia nerodnomu iazyku*. M.: Russki iasyk, 1989. *Psicologia do ensino-aprendizagem de uma língua estrangeira*. Moscou: Ed. Russki Iasyk. 1989.

ZINDER, L.R. *Obsbchaia fonetika*. M.: Vyshaia shkola, 1979. *Fonética Geral*. Moscou: Ed. Vyschaia Skola. 1979.

# **A identificação fonêmica do discurso de estudantes brasileiros na língua russa**

Tanira Castro  
Universidade Federal do Rio Grande do Sul

## **Abstract**

The Portuguese language, in the discourse of Brazilians, has specific features that are manifest in a more evident form in its lexis and phonetics, which allows us to speak, therefore, of a Brazilian Portuguese. Leaving aside the discussion of the status of Brazilian Portuguese, we must take into consideration the indisputably broad variability of its phonetic system, which is due not only to linguistic but also to sociolinguistic factors, which, in their turn, cannot but have an influence on the phonetic formation of Brazilian discourse during, for example, the learning of the Russian language (RL). Besides, the written form of some letters, both in Portuguese and Russian, causes not only phonetic but also graphic-phonetic interference to arise whether during reading, writing, or oral

discourse. The specifics of the native language of Brazilians is transferred to the discourse in the RL, mainly during the initial stage of learning. Research on the phonetic interference of the discourse of Brazilians who study the RL, is still in the early phase, but, even so, through the experimental results already obtained, we can speak of the specific and universal features of its manifestation. A priori, these violations have an orthoepic and orthophonic character, for they have a lot to do with the differential and integral criteria of the phonemes, since these violations may occur because of the specific distributive and phonotactical relations of the languages in question. The sources of the errors can be subjectively defined by the manner in which the learner assimilates and perceives the physical and linguistic functions of both languages. Both in written and oral discourse, the action of graphic-phonetic interference is observed.

A língua portuguesa falada no Brasil possui algumas características específicas, que se manifestam, principalmente, no léxico e na fonética, o que por sua vez nos permite falar sobre o status do português brasileiro. Mas deixando de lado a questão desse status, observaremos o fato indiscutível da ampla variação de seu sistema fonético, o que se deve não apenas a fatores lingüísticos, mas a questões territoriais que não poderiam deixar de influir no aspecto fonético do discurso de brasileiros durante a aprendizagem da língua russa. Além disso, a semelhança da escrita de algumas letras tanto na língua portuguesa como na russa provocam a manifestação da interferência fonético-gráfica, no discurso escrito e oral.

No português brasileiro, assim como no português europeu tradicionalmente temos sete fonemas vocálicos tônicos orais e 19 consonantais. As vogais do português brasileiro, como na língua russa diferenciam-se em anteriores, central e posteriores e podem ser altas, médias e baixas, arredondadas ou não arredondadas, mas elas ainda se diferenciam pela oposição de fonemas abertos ou fechados.

Uma questão discutível é a do status das vogais nasalizadas. No discurso brasileiro elas podem, por exemplo, ocorrer não somente em um contexto de consoantes nasais, quando poderíamos relacionar a sua manifestação à assimilação progressiva ou regressiva, mas ocorrem também antes de vogais, assim como no fim de palavras depois de uma consoante não nasal. Por exemplo – *dente, tem, irmã, dão, vão, rã, lâ*, etc. Na literatura em geral, freqüentemente as vogais nasais são analisadas como alofones das vogais puras orais. As vogais tônicas ou átonas se diferenciam ainda pela sua duração. As últimas podem sofrer uma completa redução no fim da palavra. Como já dissemos todas as vogais mencionadas podem ser tônicas. Entretanto, em um contexto antes de alguma consoante nasal – não encontramos

as vogais abertas. As vogais em posições átonas variam em dependência de suas posições em relação a sua localização no que diz respeito à tonicidade e em relação a sua posição na palavra. Desta forma, na posição pré-tônica ou pós-tônica identificamos cinco vogais: a, e, i, o, u. Acrescentamos, ainda, que o sistema vocálico português é riquíssimo em ditongos, puros e nasais (seu número é de 34), mas o seu status fonológico é discutível. Além disso, ainda é possível a combinação de três vogais – tritongos.

Como na LR, as consoantes da LP diferenciam-se conforme o ponto de articulação (labiais, bilabiais, labiodentais, dentais, pré-palatais, médio-palatais, palatais), conforme o modo de articulação (oclusivas, fricativas, explosivas, vibrantes), podem ser surdas ou soantes, sibilantes e fricativas e ainda nasais. Mas na LP não existe o critério de consoantes duras e brandas, entretanto pode manifestar-se uma forte palatalização e, em conseqüência as consoantes dentais tornam-se africadas antes das vogais [i, e], as quais na LR são reconhecidas como africadas /c'/ ou /č'/.

Após enumerarmos algumas especificidades da língua natal dos brasileiros, analisaremos o discurso russo de brasileiros, especialmente na etapa inicial de seu aprendizado. Ao analisar a interferência fonética do discurso de brasileiros que estudam a LR, ainda se encontra em fase de averiguações de suas características, mas a partir das experiências de ensino da LR no Brasil e através dos resultados experimentais já obtidos, podemos falar a respeito de suas características específicas e universais de manifestação. Como de regra, as violações cometidas têm um caráter ortoépico e ortofônico no que diz respeito aos critérios diferenciais e integrais, assim como podem ter um caráter específico de acordo com as relações distributivas e fonotáticas das línguas em contato. Tanto no discurso escrito, como no oral manifesta-se a ação da interferência fonética gráfica. A seguir enumeramos os erros fonéticos característicos e mais freqüentes no discurso russo de brasileiros com referência às vogais e às consoantes:

## 1 – Com referência às vogais, observamos

- a não diferenciação dos fonemas /b/ - /i/, o que se explica pela inexistência de tal oposição na língua natal dos aprendizes, além da semelhança articulatória acústica dessas vogais na LR. Em todas as posições fonéticas, o fonema /b/ é percebido e realizado como o fonema /i/ da língua natal dos estudantes. Observamos que tal violação manifesta-se quando o fonema /b/ encontra-se no fim da palavra e antes de uma consoante branda;
- a violação da variação alofônica das vogais russas, tendo em vista que os alunos «ignoram» o caráter ditongal das vogais brandas russas em virtude do contexto em que se encontram – consoantes brandas;
- a percepção e a realização do fonema /e/ nas sílabas tipo C'VC e C'VC' como vogal fechada anterior média, e nas sílabas tipo CVC, CVC' – como vogal aberta anterior média, o que se explica, através da identificação semelhante dos alofones dos fonemas russos com as características físicas de dois fonemas da língua materna do aluno;
- a pronúncia nasalizada de vogais não nasais russas, quando se encontram em um contexto de consoante nasal, o que evidencia a ação da variação alofônica das vogais da LM dos brasileiros;
- a percepção e a realização do fonema /o/ como uma vogal fechada posterior média, o que corresponde ao fonema da LM do aprendiz. Isso se explica pelo fato de haver uma reinterpretação das características físicas do fonema russo, a sua ditongação, semelhante ao fonema [u] – que representa o início de sua articulação;
- a vogal /a/ é bem percebida, mas a sua realização é como um fonema central baixo, correspondente ao fonema da LM do aprendiz;
- a letra 'o' russa durante a leitura pode ser realizado como o fonema [u], tendo em vista que, na LM do aprendiz, existe esse

tipo de redução, mas na LR – não. Esse tipo de erro, freqüentemente, ocorre no fim da palavra e o mesmo manifesta-se tanto na leitura como no discurso oral;

- A letra ‘**y**’ russa durante a leitura pode ser realizada como o fonema [i] de acordo com as regras de leitura da LM do aprendiz brasileiro; e,
- as letras ‘**o**’ (exceto na posição final da palavra), ‘**e**’, ‘**a**’ russas normalmente são pronunciadas como as vogais da LM do aprendiz.

2 – No que diz respeito às consoantes, temos os seguintes erros:

- os alunos não diferenciam os fonemas /c/ - /č’/, /t’/ - /č’/, /š’:/ - /č’/, /š/ - /š’:/ isso se explica não somente porque na LM do aprendiz não existe tal tipo de oposição, mas deve-se ao fato de que existe uma grande semelhança física das características dessas consoantes na LR;
- os alunos pronunciam [s] em vez de /c/, fato este observado em palavras internacionais do tipo (*natcional’ni*, *revolbiutciia*, *tcirk* e etc.), porque, na LP, nessas palavras se pronuncia [s];
- os alunos percebem e realizam os fonemas russos /x, x’/ como uvular fricativa da LM, fato este explicado pela ausência de tal fonema na LP e a sua relativa proximidade física de suas características em ambas as línguas;
- os alunos não identificam a diferença entre as consoantes duras e brandas. Além disso, as oclusivas alveolares / t’, d’/ antes de /i, e / são sempre realizadas de acordo com as regras da LM, correspondendo às africadas surda e sonora. As sonantes brandas /l’, n’/ muitas vezes podem ser realizadas como médio-palatais, como líquida e nasal, respectivamente. Neste caso a oposição obedeceu às características da LP;
- o fonema russo /t/ pode ter diferentes variantes de realização: como uma sonante línguo-alveolar fraca ou forte, como fricativa uvular sonora ou surda, como fricativa faringalizada sonora ou surda, assim como uma consoante retroflexa. Este

- último caso é característico no fim da palavra, enquanto os outros casos podem ocorrer no início ou no meio da palavra;
- a reprodução de consoantes sonoras no fim da palavra de acordo com a sua ortografia explica-se pelo fato de o aluno desconhecer as regras das terminações dos nomes (substantivos) na LR e desconhecer as respectivas normas de pronúncia;
  - a reprodução das fricativas é através da articulação dorsal de consoantes russas que são substituídas pela articulação apical ou cacuminal dorsal. A sonante lateral líquida [ɫ] na maioria das vezes articula-se como apical;
  - as fricativas sibilantes /s, z/ articulam-se como fricativas planas;
  - as consoantes duras velarizadas articulam-se como não velarizadas;
  - a realização do fonema /č/ articula-se de forma não palatalizada, mas o fonema /c/ de forma palatalizada;
  - O nome (o substantivo) russo pode finalizar com qualquer consoante e, nesses casos, os brasileiros normalmente acrescentam a vogal [i];
  - a reprodução das oclusivas linguodentais /t, d/ em combinação com a nasal /n/ articula-se sem a explosão faucal, mas na combinação com a sonante lateral líquida /ɫ/ articula-se sem a explosão lateral;
  - a consoante russa ‘**х**’ em dependência de seu contexto pode ser articulada conforme as regras fonéticas da LP, isto é como as consoantes [z], [s], [š] ou como uma combinação das consoantes [ks];
  - as consoantes [‘**Бб**’], [‘**Пп**’] articulam-se como as oclusivas bilabiais da LP, [b, p], a consoante **Cc** – como [k] ou [s]. Por sua vez, a consoante [‘**Нн**’] no início da palavra, de acordo com as regras da LP, pode não ser pronunciada, exemplo é a palavra «*nemie*» – (mudo);
  - os brasileiros não identificam e, conseqüentemente, confundem as letras russas ‘**И и**’ e ‘**И, и,**’ , assim como o ‘sinal duro’ e

o ‘sinal brando’ durante a leitura e a escrita, tendo em vista uma certa semelhança dessas letras;

- observamos também erros de pronúncia quando o aprendiz simplesmente não pronuncia algum determinado som (segmentação de subtração) ou ao contrário, acrescenta algum som a palavra (segmentação de adição), o que se deve a interferência das regras distributivas e fonotáticas da LP.

Estas são as principais características da interferência fonética no discurso de brasileiros na etapa inicial de seu aprendizado da LR. As fontes desses erros são as semelhanças físicas de suas articulações e a semelhança de suas funções lingüísticas dos fonemas em ambas as línguas.

Observamos que nossas análises consideraram o fato de que os fonetistas brasileiros descrevem o sistema consonantal de acordo com o lugar de articulação: bilabiais, labiodentais, linguodentais, alveolares, palatais, velares.

## REFERÊNCIAS BIBLIOGRÁFICAS

BAGNO, M. *Brasil e Portugal já falam duas línguas diferentes*. Galileu, fev. 2002. Nº. 127. S. 88.

BISOL, L. Introdução *a estudos de fonologia do português brasileiro*. 3. ed. Porto Alegre: EDIPUCRS, 2001.

BONDARKO, L. V. *Fonética da Língua Russa Contemporânea: Manual didático*. São Petersburgo: Editora da Universidade de São Petersburgo, 1998.

CÂMARA Jr., J. M. *Problemas de lingüística descritiva*. Petrópolis: Vozes, 1984.

CÂMARA Jr., J. M. *Estrutura da língua portuguesa*. Petrópolis: Vozes, 1970.

CASTRO, T. *A formação de habilidades auditivas e de pronúncia para estudantes lusófonos em cursos de ensino fora do meio lingüístico natural (Curso Linguafônico Fonético para Estudo sem Mestre)*. São Petersburgo: Editora da Universidade de São Petersburgo, 1995. (Tese, Doutorado em Ciências Pedagógicas.)

HOUAISS, A. *Tentativa de descrição do sistema vocálico do português culto na área dita carioca*. Rio de Janeiro: Imprensa Nacional, 1959.

LOPEZ, B. C. *The sound pattern of Brazilian Portuguese (Cariocan dialect)*. 1979. Tese (Doutorado: PhD) – University of California, Los Angeles.

MASIP, V. *Gramática de português como língua estrangeira: fonologia, ortografia e morfossintaxe*. São Paulo: EPU, 2000.

SILVA, T. C. *Fonética e Fonologia do Português: roteiro de estudos e guia de exercícios*. São Paulo: Contexto, 1999.

## NORMAS PARA APRESENTAÇÃO DE ORIGINALS

A Revista de Estudos da Linguagem publica artigos científicos, ensaios, resenhas e debates voltados para a análise de fenômenos lingüísticos.

Os trabalhos enviados serão examinados pelo Conselho Editorial desde que se apresentem de acordo com as normas abaixo relacionadas:

- 1 - Os trabalhos devem ser enviados em 3 vias digitadas em Times New Roman ou Arial, corpo 12, em Word ou Word Perfect, acompanhados de cópia em disquete.
- 2 - Os artigos e ensaios devem ser precedidos de lauda contendo nome do autor e sigla da Instituição à qual pertence, endereço para correspondência e de um resumo de 10 linhas, em inglês ou francês, para artigos escritos em português, e em português para artigos escritos em língua estrangeira.
- 3 - As ilustrações (tabelas, fotos, gráficos, etc...) deverão ser apresentadas separadamente, em arte-final e com as respectivas legendas numeradas, indicando-se, no texto, sua localização.
- 4 - No caso de ilustrações já publicadas, mencionar a fonte.
- 5 - As notas e as referências bibliográficas devem vir após o texto e ser apresentadas de acordo com a ABNT-Nbr 6023.
- 6 - As resenhas devem ser precedidas da referência bibliográfica e seguidas do nome do resenhador, sua qualificação e vinculação institucional.
- 7 - Os originais não serão devolvidos.
- 8 - Os trabalhos devem ser encaminhados a:

REVISTA DE ESTUDOS DA LINGUAGEM  
COMISSÃO EDITORIAL  
Av. Antônio Carlos, 6627  
Faculdade de Letras - sala 3025  
UNIVERSIDADE FEDERAL DE MINAS GERAIS  
31270-901 - BELO HORIZONTE - MG - BRASIL