

# Status referencial no Português Brasileiro:

## Estudo das pistas acústicas de novo, dado e acessível

Alexandre Delfino

Faculdade de Letras

Universidade Federal de Minas Gerais

Belo Horizonte, Brazil

aldelfino@gmail.com

**Abstract** — The present study has as objective to describe which acoustic parameters are related to the prosodic marking of referential status in BP. Recent proposals in the Information Structure area suggest that new and given information are in fact ends of a continuum of possible states for discourse referents. Our results show that the relative position of the referent in the utterance has a great influence in how the prosody is manifested and that duration and average  $f_0$  are parameters related to the distinction between new referents and given and accessible ones.

**Keywords-component; information structure; referential status; accessibility; prosody; Brazilian Portuguese**

**Resumo** — O presente estudo tem como objetivo a descrição dos parâmetros acústicos relacionados à marcação prosódica do status referencial no PB. Propostas recentes sobre Estrutura Informacional sugerem que informação nova e dada são na verdade dois extremos de um *continuum* de possíveis estados para referentes no discurso. Os nossos resultados mostram que a posição relativa do referente no enunciado tem uma grande influência em como a prosódia se manifesta e que a duração e a média de  $f_0$  são parâmetros relacionados à distinção de referentes novos quando comparados a referentes dados e acessíveis.

**Palavras-chave; estrutura informacional; status referencial; acessibilidade; prosódia; Português Brasileiro**

### I. INTRODUÇÃO

Na área de estrutura informacional, sabe-se que referentes possuem um estado mental temporário baseado no seu grau de disponibilidade para os interlocutores no discurso. Esse estado mental (ou status referencial) é tradicionalmente visto como uma dicotomia entre referentes novos e dados (ou velhos) no discurso [1]. Propostas recentes [2], [3] sugerem que, na verdade, o status referencial é um contínuo cujos extremos são os status novo e dado. Chafe [4], [5] propõe uma divisão baseada no grau de ativação do referente na memória de trabalho do ouvinte. Referentes novos são 'inativos', ao passo que referentes dados já estão 'ativados'. Um terceiro estado acessível é associado aos referentes 'semi-ativados'. O status referencial é muitas vezes codificado gramaticalmente, ou

seja, há uma relação razoavelmente previsível entre o status de um referente e a sua forma.

Muito se tem observado sobre a distribuição da marcação morfossintática do status referencial nas línguas: em muitas línguas (como o Português), os artigos definidos são associados a referentes dados, e os artigos indefinidos a referentes novos. Apenas recentemente há estudos que investigam a relação entre o status de um referente no discurso e a sua marcação prosódica. Em línguas germânicas (como o alemão, o inglês e o holandês), referentes novos são geralmente marcados por uma proeminência acentual ao passo que referentes dados são geralmente "desacentuados" em posição final do enunciado, ou seja, a proeminência esperada para o referente é deslocada para um item próximo, fazendo com que o referente dado seja marcado por um contorno melódico mais achatado. Baumann [6] observou ainda que no alemão referentes acessíveis podem ser marcados por uma proeminência intermediária, de acordo com a relação semântica entre o referente e o seu antecedente, embora o autor não tenha definido uma configuração acentual (pitch accent) típica para esse status. Aparentemente, a marcação acentual do status acessível só é consistente quando a relação é de "todo-pela-parte". Em outras línguas, como o italiano, o árabe e o romeno, observou-se que referentes dados não são desacentuados em posição final. Nesse caso, a configuração acentual de itens novos e dados é a mesma [7], [8], [9]. Em Português Brasileiro, trabalhos recentes como [10], [11], [12] mostram que a prosódia é sensível à codificação do status referencial. Delfino et al. [11] analisou a marcação prosódica em três níveis (novo, dado e acessível). Os resultados mostraram que referentes novos são distintos acusticamente de referentes dados e acessíveis, mas não há diferença entre esses dois últimos. Os autores concluíram que a falta de controle da relação semântica entre os itens acessíveis e seus respectivos antecedentes pode ter influenciado nos resultados. Antão et al. [12] analisou a relação entre marcação acústica do status referencial e a posição do referente na sentença. Os resultados mostram falantes tendem a marcar de forma mais explícita o status referencial quando referentes estão na posição inicial. Esse estudo, contudo, considerou apenas os status novo e dado.

O objetivo do presente trabalho, portanto, é aprofundar-se em algumas questões relacionadas à manifestação prosódica do status referencial, considerando principalmente dois pontos: (a) a posição relativa do referente no enunciado e (b) a relação semântica do status acessível.

## II. METODOLOGIA

Para este experimento, foram criados dois corpora com 72 narrativas, distribuídas em três condições: dado, novo e acessível. O contexto das sentenças determinava o status referencial da palavra-alvo. No primeiro corpus, as palavras-alvo estão em posição final de palavra, e no segundo corpus as palavras-alvo estão em posição inicial não absoluta. As sentenças abaixo ilustram a construção do status referencial pelo contexto (as palavras-alvo estão em **negrito** e os antecedentes em *itálico*):

TABELA I. EXEMPLOS DE NARRATIVAS UTILIZADAS NO EXPERIMENTO

	Posição inicial
Novo	De acordo com o manual, <b>uma bateria</b> é o suficiente para ligar o brinquedo novo.
Dado	O rapaz comprou <i>uma bateria</i> para o aparelho. Depois de alguns minutos, <b>a bateria</b> apresentou um cheiro de queimado.
Acessível	Ao sair de casa, o rapaz não conseguiu ligar <i>o carro</i> . Durante a madrugada, <b>a bateria</b> havia congelado.
	Posição final
Novo	O homem foi à loja para comprar <b>uma bateria</b> . O rapaz comprou <i>uma bateria</i> nova para o seu notebook.
Dado	Quando ele sentiu um cheiro de queimado, ele foi à loja devolver <b>a bateria</b> .
Acessível	Ao sair de casa, o rapaz não conseguiu ligar <i>o carro</i> . O frio havia congelado <b>a bateria</b> .

Todas as palavras-alvo possuem quatro sílabas e são paroxítonas. Arantes [10] mostrou que os correlatos acústicos para a marcação do status referencial se tornam mais evidentes à medida que o número de sílabas da palavra aumenta. A escolha por apenas palavras paroxítonas possui duas razões: primeiro, ela representa o padrão acentual mais comum na língua e, segundo, ela permite o controle de alguma proeminência acentual durante a análise acústica. A relação semântica entre o antecedente e a palavra-alvo na condição acessível foi controlada: todos os referentes acessíveis possuem uma relação “todo-pela-parte” (por exemplo, *carro~bateria*) com seus antecedentes. O traço 'animacidade' também foi controlado em todos os itens.

O experimento foi realizado em duas etapas, utilizando-se um corpus diferente por vez. Em cada etapa, nove participantes (quatro para a primeira etapa e cinco para a segunda) tiveram como tarefa a leitura da lista de sentenças, que foram apresentadas uma a uma em uma apresentação de slides. Os

participantes controlavam a velocidade de transição entre os slides. Eles foram instruídos a realizar uma leitura silenciosa antes da gravação de cada sentença. Durante a análise acústica, foram considerados os seguintes parâmetros: (a) duração da palavra-alvo, (b) média da curva de  $f_0$ , (c) desvio-padrão, (d) gama e (e) contorno de  $f_0$  normalizado pelo tempo. Para essas medidas, foram utilizados scripts do Praat que extraíram os valores e calcularam o valor médio por cada parâmetro. Os valores para (a) foram calculados em milissegundos, para (b) e (c) em Hetz e para (d) em semitons. Este último foi obtido aplicando-se a fórmula abaixo, em que  $f_{0max}$  e  $f_{0min}$  são respectivamente os valores de  $f_0$  máximo e mínimo do contorno melódico:

$$12 \log_2(\max_{Hz}/\min_{Hz})$$

Os dados dos participantes foram analisados separadamente. O status referencial com três níveis (dado, novo e acessível) era a variável independente. Para a análise estatística, foram feitas Análises de Variância (ANOVAs) para determinar se as diferenças nos valores médios dos parâmetros (a)-(d) eram estatisticamente diferentes. O teste de Bonferroni (post-hoc) foram realizados para ajustar o nível de alfa para comparações múltiplas. Para a análise de (e), os contornos entonacionais foram normalizados pelo tempo, seguindo um script proposto em [10]. Os contornos foram suavizados e interpolados. Por fim, foram extraídas cinco amostras de  $f_0$  para cada sílaba e o artigo de cada palavra-alvo. A análise de (e) permitiu uma observação visual do comportamento da curva de  $f_0$  como um todo, considerando-se a palavra-alvo e o artigo.

## III. RESULTADOS

Os resultados da primeira etapa mostram que o status referencial afeta de forma significativa a duração e a média de  $f_0$  nas palavras-alvo na maioria dos participantes. Nos outros correlatos de  $f_0$  (desvio-padrão e gama), não houve diferença significativa entre as condições. Por outro lado, os resultados da segunda etapa mostram que não houve qualquer diferença entre as três condições na posição final da sentença.

A análise do contorno entonacional normalizado pelo tempo mostra que, em posição inicial, referentes novos podem ser caracterizados por dois picos no contorno, um alinhado às sílabas pretônicas e outro alinhado à sílaba tônica. Referentes dados e acessíveis mostram contornos muito similares e tendem a ter apenas um pico alinhado à sílaba tônica e um contorno relativamente achatado (Figura 1). Em posição final, o contorno entonacional em todas as condições é muito similar, com apenas um pico alinhado à sílaba tônica (Figura 2).

## IV. DISCUSSÃO

De forma geral, os dados do experimento corroboram os resultados encontrados em outros estudos sobre a marcação prosódica do status referencial no PB. Aparentemente, a

posição do referente no enunciado parece afetar a forma como os falantes utilizam os diferentes parâmetros acústicos. Na primeira etapa (itens-alvo na posição inicial), as diferenças na duração e no valor médio de pitch se mostram significativos para marcar o status referencial. Entretanto, quando os itens-alvo estavam na posição final, nenhum dos correlatos acústicos mostrou qualquer diferença significativa.

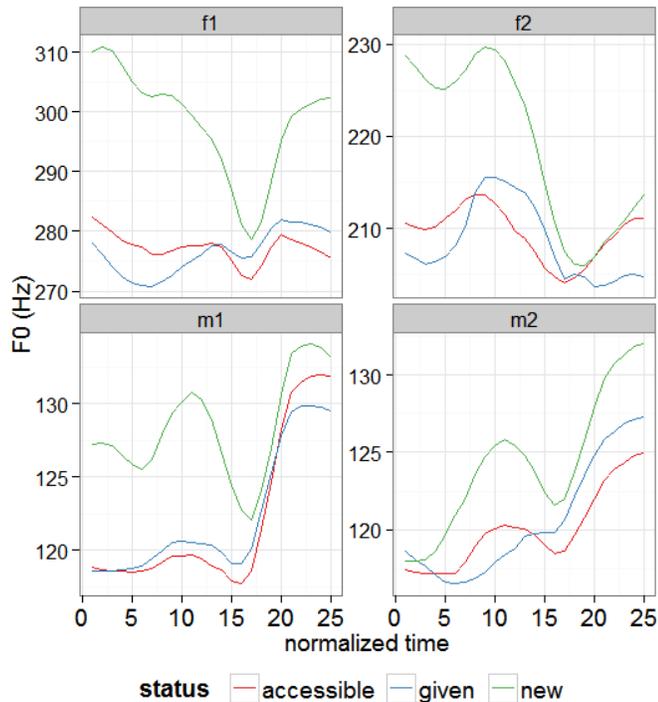


Figura 1. Contornos de f0 normalizados para referentes na posição inicial. Barras verticais indicam limites entre sílabas.

Uma possível explicação para os resultados encontrados vem de propostas que sugerem uma ordem de distribuição dos referentes na estrutura informacional [1], [13], [14]. Gundel [13], através do Princípio Dado-antes-de-novo, propõe que o enunciado segue um princípio em que informação dada tende a ocorrer no início do enunciado e informação nova tende a ocorrer nas posições finais. O Modelo de Syntaxe-Discurso [14], de forma similar, propõe que a informação no início do enunciado tende a ser dada, pois ela serve como um elo para outras unidades discursivas já mencionadas. Se a posição final é *par excellence* para informação nova, então uma proeminência acústica para marcar referentes novos parece menos necessária. Nesse caso, os ouvintes poderiam utilizar de outros tipos de informação linguística para marcar o status referencial (informação morfossintática, por exemplo).

De forma análoga, quando referentes novos ocorrem em posição inicial, os falantes necessitam sinalizar aos ouvintes de que aquela informação não foi previamente mencionada no discurso. Essa proposta é corroborada pelo Effort Code [15], [16]. Falantes de uma língua compartilham um código biológico gramaticalizado em que uma informação relevante é

marcada com um esforço articulatorio maior. O falante deve indicar que há uma informação relevante onde se esperaria uma informação não relevante. Dessa forma, a ocorrência de informação não mencionada em posição inicial é marcada e deve receber uma proeminência acústica de forma a sinalizar que ela de fato é nova (e, portanto, relevante). No caso do PB, essa sinalização é feita com um aumento da duração e do nível médio de pitch.

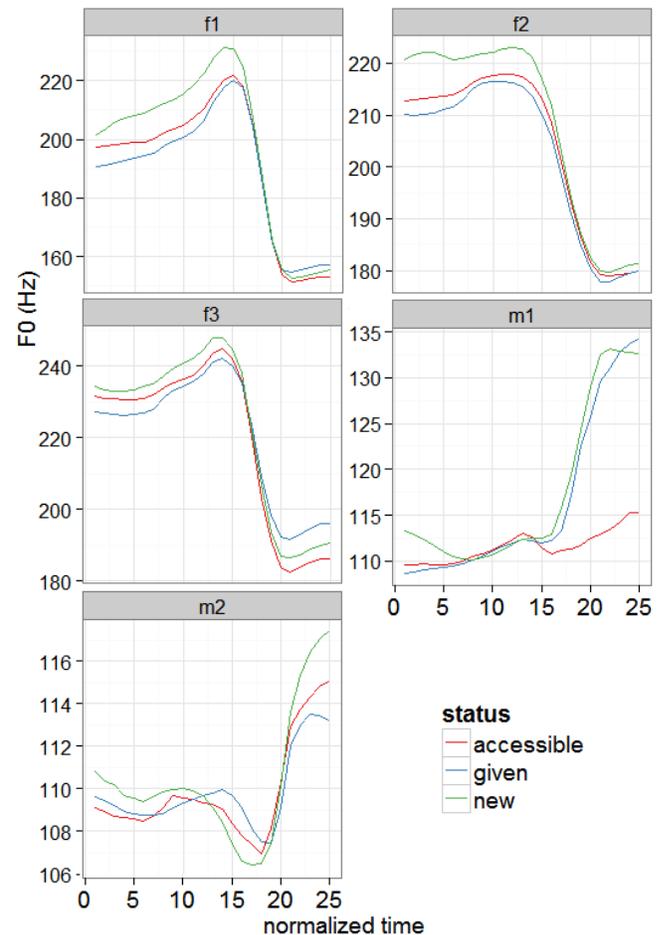


Figura 2. Contornos de f0 normalizados para referentes na posição final.

A análise dos referentes acessíveis indica que, ao contrário do que foi encontrado no alemão [6], eles não possuem nenhum correlato acústico comparado aos referentes dados. Aparentemente, referentes acessíveis se comportam de forma similar aos referentes dados, o que possibilita alguns questionamentos.

Primeiramente, é possível que, durante o experimento, os sujeitos fossem guiados por pistas morfossintáticas (artigos definidos e indefinidos) e não pela informação contextual. Como os referentes acessíveis e dados são marcados com artigos definidos, isso poderia ter levado a uma identificação dos itens-alvo como um único status. Uma evidência para essa interpretação é que não houve diferença significativa em

nenhum dos parâmetros acústicos analisados. Em segundo lugar, como Baumann [6] apontou, a marcação acústica de referentes acessíveis é variável, e que em apenas um caso (quando referentes e antecedentes possuem uma relação de todo-pela-parte) houve uma marcação típica do tipo H+L\*. Há ainda outras propostas que negam um correlato fonológico direto para informação acessível [17]. Chafe [5] ainda propõe que referentes acessíveis, estando em um estado semi-ativado, teriam uma proeminência prosódica similar a referentes novos. A nossa análise empírica mostra, contudo, que referentes acessíveis possuem uma codificação prosódica similar a referentes dados. Finalmente, devemos considerar a possibilidade de que a marcação de status intermediários no continuum de status referenciais é variável entre as línguas. Conforme vimos na Introdução, as línguas podem codificar prosodicamente o status referencial de formas diferentes. Enquanto as línguas germânicas alteram a configuração entonacional de referentes dados (desacentuação), as línguas românicas tendem a manter a configuração (reacentuação). Nesse cenário, diferenças tipológicas entre as línguas poderiam explicar porquê referentes acessíveis em PB tendem a se comportar como referentes dados e em alemão eles possuem uma marcação acentual proeminente. De toda forma, é necessário que haja mais estudos exploratórios sobre essa questão. Até o presente momento, os nossos dados são comparáveis apenas aos dados do alemão [6]. Trabalhos futuros poderão considerar as hipóteses discutidas para um entendimento da marcação acústica do status acessível em maior profundidade.

Por fim, outra questão que chamou a atenção foi a variabilidade individual. Ao longo do experimento, os sujeitos variaram a combinação de média do nível de *pitch* e duração; em alguns sujeitos, a gama mostrou-se com parâmetro relevante. Como um fenômeno pragmático, assumimos que outros fatores podem também estar relacionados à variação entre os sujeitos, como registro, estilo pessoal, foco enfático e dialeto.

## V. CONCLUSÃO

Esse estudo teve como objetivo investigar quais correlatos acústicos estão relacionados à marcação do status referencial no PB, considerando-se a posição relativa dos referentes no enunciado. Os resultados corroboram estudos anteriores [10], [11], [12] e mostram que a informação prosódica é afetada pelo status referencial no discurso. A posição dos referentes afetou a forma como os referentes são marcados acusticamente. Na posição inicial, a duração e a média de *f<sub>0</sub>* marcam uma distinção entre o status novo e os status dado e acessível. Na posição final, os status não possuem uma marcação distintiva. A nossa hipótese é de que essa diferença é consequência da organização geral da estrutura informacional do enunciado, em que informação nova é esperada em posição final e portanto não necessita de uma

marcação distintiva. Por fim, a falta de diferença acústica entre os status dado e acessível permite uma série de possibilidades, que vai desde uma influência de outros tipos de informação (por exemplo, morfossintática) até diferenças tipológicas entre as línguas.

## AGRADECIMENTOS

O autor gostaria de agradecer a Maria Luiza Cunha Lima, Petra B. Schumacher e Stefan Baumann pelas sugestões para o desenho experimental e os materiais utilizados e a Pablo Arantes pelo apoio na análise estatística e na produção dos gráficos.

## REFERÊNCIAS

- [1] M. A. Halliday, "Notes on transitivity and theme in English: Part 2," *J. of Linguistics*, vol. 3, no 2, pp. 199-244, 1967.
- [2] Prince, E. 1981. "Towards a taxonomy of given-new information." In Cole, P. *Radical Pragmatics*, New York, Academic Press: 223-256.
- [3] J. K. Gundel, N. Hedberg, & R. Zacharski, "Cognitive status and the form of referring expressions in discourse," *Language*, vol 69, no 2, pp. 274-307, 1993.
- [4] W. L. Chafe, "Givenness, contrastiveness, definiteness, subjects, topics and point of view," in *Subject and topic*, N. Li Charles, Ed. New York: Academic Press, 1976, pp. 25-55.
- [5] W. L. Chafe, *Discourse, consciousness, and time*. Chicago: The University of Chicago Press, 1994.
- [6] S. Baumann, "The Intonation of Givenness - Evidence from German," *Linguistische Arbeiten 508*. Tübingen: Niemeyer, 2006.
- [7] M. Swerts, E. Kraemer, and C. Avesani, "Prosodic marking of information status in Dutch and Italian: A comparative analysis," *J. of Phonetics*, vol. 30, no. 4, pp. 629-654, 2002.
- [8] S. Hellmuth, "No De-accenting in (or of) Phrases: Evidence from Arabic for cross-linguistic and cross-dialectal prosodic variation," *Prosodies: with special reference to Iberian languages*, S. Frota, M. Vigário, M.J. Freitas, Ed. Walter de Gruyter, 2005, pp. 99-112.
- [9] M. Swerts, "Contrast and accent in Dutch and Romanian," *J. of Phonetics*, vol. 35, no. 3, pp. 380-397, 2007.
- [10] P. Arantes, "Integrando produção e percepção de proeminências numa abordagem dinâmica do ritmo da fala," Ph.D. dissertation. Unicamp, Campinas, 2010.
- [11] A. Delfino, P. Arantes and M. L. Cunha Lima, "Prosodic marking of referential status in Brazilian Portuguese: a preliminary study," in *Proc. GSCP 2012 Intern. Conf*, Belo Horizonte, 2012, pp. 186-190.
- [12] C. Antão, M. L. Cunha Lima, and P. Arantes, "Interrelation between subjecthood, referential status and prosody," in *Proc. 35th Annual Conf. German Linguistic Society – DGfS*. Potsdam, 2013, pp. 55-56.
- [13] J. K. Gundel, *The role of topic and comment in linguistic theory*, Garland Pub, 1988.
- [14] P. Burkhardt, *The syntax-discourse interface: Representing and interpreting dependency* vol. 80, John Benjamins Publishing, 2005.
- [15] C. Gussenhoven, "Intonation and interpretation: phonetics and phonology," presented at the Speech Prosody 2002, Aix-en-Provence, 2002.
- [16] C. Gussenhoven, *The phonology of tone and intonation*, Cambridge University Press, 2004.
- [17] K. Lambrecht, *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*, Cambridge University Press, 1994.