

Maria Mendes Cantoni

**O ACENTO NO PORTUGUÊS BRASILEIRO:**  
**uma abordagem experimental**

Maria Mendes Cantoni

# **O ACENTO NO PORTUGUÊS BRASILEIRO:** **uma abordagem experimental**

Tese apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Linguística Teórica e Descritiva.

Área de concentração: Linguística Teórica e Descritiva

Linha de Pesquisa: Fonologia

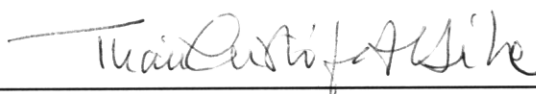
Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Thaís Cristófaró Silva

Co-orientador: Prof. Dr. Rafael Laboissière

Belo Horizonte  
Faculdade de Letras da UFMG

2013

Tese intitulada *O acento no português brasileiro: uma abordagem experimental*, defendida por MARIA MENDES CANTONI em 25/01/2013 e aprovada pela Banca Examinadora constituída pelos Professores:



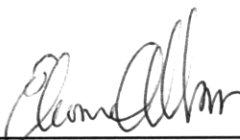
---

**Dra. Thaís Cristóvão Alves da Silva - UFMG**  
**Orientadora**



---

**Dr. Rafael Michelin Laboissière - Centre National de la Recherche Scientifique-  
França**



---

**Dra. Eleonora Cavalcante Albano - UNICAMP**



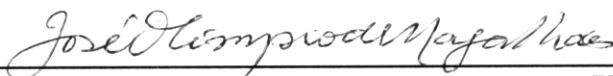
---

**Dr. Pablo Arantes - UFRJ**



---

**Dr. Maurílio Nunes Vieira - UFMG**



---

**Dr. José Olímpio de Magalhães - UFMG**

## Agradecimentos

Este trabalho só foi possível graças à contribuição e apoio de diversas pessoas.

Agradeço especialmente à Profa. Thaïs Cristófaró Silva, pela orientação e apoio de valor inestimável ao longo dos últimos nove anos.

Ao Prof. Rafael Laboissière, co-orientador deste trabalho, pela gentil acolhida em seu laboratório em Lyon, pela análise estatística dos dados, pela disponibilização do Morphala para geração de estímulos e pela paciência e generosidade.

Ao Prof. Maurílio Vieira, pela cessão de equipamentos e programa de análise do EGG, pelo auxílio metodológico e pela disponibilidade e gentileza. Agradeço também pelos comentários e sugestões no exame de qualificação e na defesa desta tese.

Ao Prof. Hani Yehia e sua equipe do Cefala, pela cessão de estrutura física, equipamentos, apoio técnico e auxílio na análise dos dados.

Ao Prof. César Reis, pelos comentários e sugestões no exame de qualificação.

À Profa. Sandra Bianchet, por comentários no capítulos a respeito do latim.

Ao Prof. José Olímpio de Magalhaães, pela avaliação do projeto de pesquisa e comentários na defesa desta tese.

À Profa. Eleonora Albano, pelos comentários na defesa desta tese.

A Pablo Arantes, pelo auxílio com as ferramentas de medição acústica e pelos comentários na defesa desta tese.

A João Pedro Sansão, José Eduardo Silva e Leonardo Almeida, por toda ajuda na coleta e pré-processamento dos dados.

Aos colegas do laboratório do Elabore, pelo suporte, companheirismo e tão preciosos momentos de descontração.

À Capes e ao CNPq, pelo suporte financeiro durante os quatro anos de doutorado e os quatro meses de estágio de doutoramento “sanduíche” em Lyon, França.

A Janes, minha mãe, pela revisão criteriosa do texto. Agradeço também a ela e a Rodrigo, meu esposo, por todo o apoio e compreensão ao longo dos últimos quatro anos.

## Resumo

Este estudo trata do acento lexical no português brasileiro (PB) sob uma perspectiva multirrepresentacional. Tem como principal objetivo avaliar mecanismos acústicos, articulatórios e cognitivos envolvidos na produção e percepção do acento. A metodologia adotada conjuga a análise teórica fonológica com evidências experimentais, seguindo a proposta da Fonologia de Laboratório. Foram realizados um experimento de produção e três experimentos de percepção. No experimento de produção, foram avaliadas propriedades acústicas e articulatórias envolvidas na realização do acento, a partir da gravação dos sinais de fala e eletroglotográfico. Verificou-se que a principal propriedade acústica empregada na realização do acento no PB é a duração, capaz de diferenciar sílabas acentuadas de pretônicas e postônicas e as últimas entre si. Na série de experimentos de percepção, foi avaliada a interação entre acento e acesso lexical. O primeiro experimento testou a influência da frequência de tipo na segmentação de sequências não linguísticas ambíguas com relação à localização da proeminência. Os resultados do primeiro experimento indicam que as propriedades acústicas exercem influência na percepção da proeminência em padrões sonoros. O segundo experimento de percepção testou a influência da frequência lexical na segmentação de sequências linguísticas ambíguas com relação à posição da proeminência. Finalmente, o terceiro experimento testou a influência da frequência lexical na percepção categórica do acento, a partir de estímulos gerados por *morphing*. O segundo e terceiro experimentos mostraram que a percepção da proeminência acentual depende da frequência de ocorrência das palavras. Com base nos resultados, argumenta-se que informações sobre o acento são armazenadas lexicalmente, o que explicaria com sucesso os efeitos de frequência encontrados nos experimentos de percepção. A proeminência no nível da palavra promovida pelo acento é interpretada como resultado de generalizações linguísticas e cognitivas, generalizações estas responsáveis pelas tendências gerais na distribuição do acento no léxico e sua interação com a gramática. É proposta uma análise em redes do fenômeno do acento, que incorpora princípios de funcionamento de sistemas dinâmicos. Nessa proposta, a acentuação do português teria emergido pela auto-organização de variáveis do sistema acentual latino em decorrência de uma série de mudanças fonológicas desencadeadas pela perda da quantidade vocálica. Considerando fenômenos sonoros do PB atual, a presença de mais de uma pista acústica para a sílaba acentuada (relação de troca) possibilitou que duração passasse a exercer um papel relevante no contraste entre formas que sofreram redução da vogal postônica e formas plenas correspondentes.

Palavras-chave: Acento; Modelo de Uso; Fonologia de Laboratório; Prosódia; Eletroglotografia; Percepção categórica

## Abstract

This thesis surveys the lexical stress in Brazilian Portuguese (BP) according to a multi-representational perspective. The main goal of this study is to evaluate acoustic, articulatory and cognitive mechanisms involved in stress production and perception. In order to refine the phonological analysis with empirical evidences, a total of four experiments were designed: one regarding the production of stress and the other three regarding the perception of stress. The production experiment was designed to evaluate acoustic and articulatory properties. We recorded speech and electroglottographic signal. The results of this experiment suggest that the main acoustic property of stress in BP is duration, which can faithfully differentiate stressed syllables from pretonic, stressed syllables from posttonic as well as pretonic from posttonic. In the series of the three perception experiments, we evaluated the interaction between stress and lexical access. The first experiment tested the parsing of ambiguous non-linguistic sequences with regard to the location of the prominence. The results from the first experiment indicate that acoustic properties influence the perception of prominence of sound patterns. The second experiment tested the influence of lexical frequency on the parsing of ambiguous linguistic sequences with regard to the location of the prominence. The third experiment tested the influence of lexical frequency on the categorical perception of stress, using stimuli generated by *morphing*. The second and third experiments show that the perception of stress prominence in linguistic sound patterns depends on the token frequency of the words. The findings support the idea that information on stress is lexically stored, which can successfully account for the frequency effects found in the previous perception experiments. Additionally, such findings can be supported by the that the word-level prominence involved in lexical stress can be interpreted as a result of linguistic and cognitive generalizations, which account for general tendencies in the distribution of stress in the lexicon, besides its interaction with grammar. We analyze stress as an emergent abstract pattern within a network model framework, which incorporates working principles of dynamic systems. In this analysis, we state that stress in BP might have emerged from the self-organization variables found in the stress system of Latin, as a consequence of a chain of phonological changes after vowel quantity loss. Regarding present-day BP phonological phenomena, we further argue that the multiplicity of acoustic cues in stressed syllables (trading relationship) had duration play a relevant role in contrasting reduced forms of posttonic vowels with their non-reduced counterparts.

Key-words: Stress; Usage-Based Model; Laboratory Phonology; Prosody; Electroglottography; Categorical perception

## Sumário

<b>1 Introdução.....</b>	<b>11</b>
<b>2 Sobre o estudo científico da linguagem .....</b>	<b>14</b>
2.1 Primórdios da linguística.....	14
2.2 Modelagem formal: a linguagem como sistema serialista computacional .....	16
2.3 Surgimento da Linguística Cognitiva.....	18
2.4 Abordagens multirrepresentacionais.....	26
<b>3 Acento e produção de fala.....</b>	<b>36</b>
3.1 Delimitação do fenômeno acentual.....	36
3.2 O estudo do acento ao longo do desenvolvimento histórico das teorias fonológicas...39	
3.2.1 Estruturalismo.....	39
3.2.2 Gerativismo e modelos não lineares.....	43
3.2.3 Modelos gestuais.....	45
3.3 Aspectos físicos da proeminência acentual.....	46
3.4 Estudos sobre o acento lexical no português.....	48
<b>4 Percepção.....</b>	<b>57</b>
4.1 A complexidade da percepção .....	57
4.2 Percepção categórica.....	59
4.3 Percepção e aspectos prosódicos.....	62
4.4 Perspectivas teóricas.....	63
<b>5 Abordagem multirrepresentacional do acento.....</b>	<b>65</b>
5.1 Percurso histórico do acento lexical português.....	65
5.2 Abordagem multirrepresentacional.....	74
<b>6 Análise experimental do acento.....</b>	<b>84</b>
6.1 A metodologia experimental no estudo do acento.....	84
6.2 Experimento de produção.....	85
6.2.1 Sujeitos.....	86
6.2.2 Procedimentos.....	86
6.2.3 Estímulos.....	88
6.2.4 Pré-processamento dos dados.....	92
6.2.5 Medição.....	96
6.2.6 Resultados.....	100
6.2.7 Discussão dos resultados.....	120
6.2.8 Conclusões.....	129

6.3 Experimentos de percepção.....	133
6.3.1 Experimento 1: desambiguação de padrões não linguísticos.....	134
6.3.1.1 Sujeitos.....	135
6.3.1.2 Estímulos.....	136
6.3.1.3 Procedimentos.....	136
6.3.1.4 Resultados.....	137
6.3.1.5 Discussão dos resultados.....	141
6.3.2 Experimento 2: desambiguação de padrões linguísticos.....	144
6.3.2.1 Sujeitos.....	144
6.3.2.2 Estímulos.....	144
6.3.2.3 Procedimentos.....	147
6.3.2.4 Resultados.....	147
6.3.2.5 Discussão dos resultados.....	150
6.3.3 Experimento 3: percepção categórica.....	152
6.3.3.1 Sujeitos.....	153
6.3.3.2 Estímulos.....	153
6.3.3.3 Procedimentos.....	158
6.3.3.4 Resultados.....	159
6.3.3.5 Discussão dos resultados.....	164
<b>7 Conclusões.....</b>	<b>167</b>
<b>Apêndice A.....</b>	<b>170</b>
<b>Apêndice B.....</b>	<b>171</b>
<b>Apêndice C.....</b>	<b>172</b>
<b>Referências.....</b>	<b>182</b>

## Lista de Ilustrações e Tabelas

FIGURA 1 - Comparação entre modelos linguísticos tradicionais e multirrepresentacionais. .27	27
FIGURA 2 - Estrutura simbólica da linguagem segundo a Linguística Cognitiva .....29	29
FIGURA 3 - Rede de exemplares e emergência de padrão morfológico.....35	35
FIGURA 4 - Medidas básicas realizadas em uma onda.....47	47
FIGURA 5 - Média de duração e duração relativa em três variedades do português.....54	54
FIGURA 6 - Propriedades acústicas do acento lexical em início de sentença.....55	55
FIGURA 7 - Propriedades acústicas do acento lexical em final de sentença.....56	56
FIGURA 8 - Exemplo de curva psicométrica obtida em experimento de identificação.....62	62
FIGURA 9 - Representação de redes e esquemas acentuais no latim.....68	68
FIGURA 10 - Representação de redes morfológicas e acentuais no latim.....70	70
FIGURA 11 - Níveis prosódicos na metrificação latina.....73	73
FIGURA 12 - Estrutura prosódica emergente e esquemas de nível alto para proparoxítonas, paroxítonas e oxítonas em verbos.....79	79
FIGURA 13 - Estrutura prosódica emergente e esquemas de nível alto para proparoxítonas, paroxítonas e oxítonas em não verbos.....80	80
FIGURA 14 - Projeto experimental utilizado na coleta dos dados de produção .....88	88
FIGURA 15 - Estrutura de camadas utilizada na segmentação do sinal obtido no experimento de produção.....94	94
FIGURA 16 - Pontos de referência utilizados nas medidas da articulação glótica a partir do sinal do EGG .....99	99
FIGURA 17 - Exemplos de dados obtidos no experimento de produção .....102	102
FIGURA 18 - Distribuição dos dados de duração (s) de acordo com tipo de proeminência, localização da proeminência e foco.....106	106
FIGURA 19 - Distribuição dos dados de F0 (Hz) de acordo com tipo de proeminência, localização da proeminência e foco.....108	108
FIGURA 20 - Distribuição dos dados de intensidade (dB) de acordo com tipo de proeminência, localização da proeminência e foco.....110	110
FIGURA 21 - Distribuição dos dados de quociente de contato CQ (%) de acordo com tipo de proeminência, localização da proeminência e foco.....112	112

FIGURA 22 - Distribuição dos dados de índice de velocidade SI (u.a.) de acordo com tipo de proeminência, localização da proeminência e foco .....	114
FIGURA 23 - Distribuição dos dados de fase de fechamento Cp (%) de acordo com tipo de proeminência, localização da proeminência e foco.....	116
FIGURA 24 - Distribuição dos dados de fase de abertura Op (%) de acordo com tipo de proeminência, localização da proeminência e foco.....	118
FIGURA 25 - Propriedades acústicas do acento obtidas experimentalmente.....	123
FIGURA 26 - Projeto experimental utilizado nos três experimentos de percepção.....	134
FIGURA 27 - Sequências de tons utilizadas como estímulos no experimento de percepção 1 .....	137
FIGURA 28 - Estímulos visuais utilizados no experimento de percepção 1.....	138
FIGURA 29 - Distribuição das probabilidades de resposta de acordo com o padrão sonoro dos estímulos.....	140
FIGURA 30 - Exemplo de estímulo utilizado no experimento de percepção 2.....	147
FIGURA 31 - Percentual de respostas paroxítonas em função da frequência de ocorrência relativa das paroxítonas .....	150
FIGURA 32 - Diagrama esquemático do processo de construção dos estímulos do experimento de percepção 3.....	155
FIGURA 33a - Exemplos de estímulos gerados pelo morphing no experimento de percepção 3 .....	156
FIGURA 33b - Exemplos de estímulos gerados pelo morphing no experimento de percepção 3 .....	157
FIGURA 33c - Exemplos de estímulos gerados pelo morphing no experimento de percepção 3 .....	158
FIGURA 34 - Valores das respostas a cada par de estímulos no extremo 0%.....	162
FIGURA 35 - Valores das respostas a cada par de estímulos no extremo 100% .....	163
FIGURA 36 - Correlação entre intercepto dos conjuntos de estímulos e frequência de ocorrência relativa da paroxítona ( $R = 0,78$ , $t[7] = 3,28$ , $p < 0,05$ ).....	164
FIGURA 37 - Inclinação vs. intercepto do fator aleatório sujeito.....	165
FIGURA 38 - Curva psicométrica do conjunto de estímulos come-comi.....	174
FIGURA 39 - Curva psicométrica do conjunto de estímulos fabrica-fábrica.....	175
FIGURA 40 - Curva psicométrica do conjunto de estímulos ficaram-ficarão.....	176

FIGURA 41 - Curva psicométrica do conjunto de estímulos hesito-êxito.....	177
FIGURA 42 - Curva psicométrica do conjunto de estímulos pito-pitu.....	178
FIGURA 43 - Curva psicométrica do conjunto de estímulos pratica-prática.....	179
FIGURA 44 - Curva psicométrica do conjunto de estímulos publica-pública.....	180
FIGURA 45 - Curva psicométrica do conjunto de estímulos secretária-secretaria.....	181
FIGURA 46 - Curva psicométrica do conjunto de estímulos tato-tatu.....	182
QUADRO 1 - Sistema de traços fonológicos proposto em Jakobson e Halle (1971).....	42
QUADRO 2 - Aspectos a serem abordados na análise multirrepresentacional do acento proposta neste estudo.....	57
QUADRO 3 - Sílabas testadas e seu contexto adjacente conforme as variáveis sílabas e tipo	91
QUADRO 4 - Códigos utilizados para anotação do sinal estéreo.....	95
QUADRO 5 - Critérios adotados na segmentação do sinal de fala .....	96
QUADRO 6 - Resumo dos testes simultâneos de comparações múltiplas de médias.....	120
QUADRO 7 - Tendências de favorecimento e desfavorecimento à percepção padrões .....	143
QUADRO 8 - Pares de estímulos utilizados no experimento de percepção 2 e seus valores de ocorrência.....	146
QUADRO 9 - Palavras associadas aos conjuntos de estímulos utilizados no experimento de percepção 3 e seus valores de ocorrência.....	159
TABELA 1 - Frequência de tipo de cada padrão acentual em verbos e não verbos.....	83
TABELA 2 - Frequência de ocorrência de cada padrão acentual em verbos e não verbos.....	83
TABELA 3 - Resultados dos testes de comparações múltiplas para a variável duração.....	107
TABELA 4 - Resultados dos testes de comparações múltiplas para a variável F0.....	109
TABELA 5 - Resultados dos testes de comparações múltiplas para a variável intensidade...	111
TABELA 6 - Resultados dos testes de comparações múltiplas para a variável CQ.....	113
TABELA 7 - Resultados dos testes de comparações múltiplas para a variável SI.....	115
TABELA 8 - Resultados dos testes de comparações múltiplas para a variável Cp.....	117
TABELA 9 - Resultados dos testes de comparações múltiplas para a variável Op.....	119
TABELA 10 - Valores médios de CQ (%).....	126
TABELA 11 - Valores médios de SI (u.a.).....	127
TABELA 12 - Valores médios de Cp (%).....	128
TABELA 13 - Valores médios de Op (%).....	129

## 1 Introdução

O presente estudo tem como tema nuclear o acento lexical no português brasileiro (doravante PB), ou seja, a proeminência conferida a uma das sílabas de cada palavra nesta língua. Considerando o acento como um fenômeno que acontece na dimensão acústica, articulatória e cognitiva, pretende-se refletir sobre os mecanismos responsáveis por sua produção e percepção. As três dimensões do acento serão abordadas seguindo a proposta da Fonologia de Laboratório (PIERREHUMBERT; BECKMAN; LADD, 2000), conjugando uma análise teórica fonológica com evidências experimentais a respeito de (1) propriedades acústicas e articulatórias envolvidas na produção do acento; (2) interação entre proeminência acentual e acesso lexical na percepção do acento.

Nesse contexto, esta tese tem como objetivo geral analisar os mecanismos de atribuição do acento lexical no português e, dessa forma, contribuir para a construção de conhecimento científico acerca da dinâmica de processamento e armazenamento de informações sonoras e prosódicas.

Quanto aos objetivos específicos, pretende-se avaliar:

- a) parâmetros acústicos e articulatórios envolvidos na produção da proeminência acentual;
- b) parâmetros acústicos relevantes para a percepção da proeminência acentual;
- c) a natureza representacional dos mecanismos de atribuição de acento;
- d) o impacto da frequência de uso na construção das representações mentais;
- e) o papel de aspectos prosódicos, morfológicos e segmentais, bem como sua interação, para a dinâmica acentual no PB.

A argumentação a ser desenvolvida nos próximos capítulos buscará demonstrar que o acento no PB é armazenado no léxico, o que está em consonância com a ideia de que informações redundantes estão presentes na representação mental (JOHNSON; MULLENIX, 1997; PIERREHUMBERT, 2001a). A acentuação como mecanismo promotor de proeminência no nível da palavra será interpretada como resultado de generalizações linguísticas e cognitivas que são responsáveis pelas tendências gerais na distribuição do acento no léxico e sua interação com a gramática. Esta tese apresenta como proposta uma modelagem em redes do fenômeno do acento, que incorpora princípios de funcionamento dos

sistemas dinâmicos (BECKNER et al., 2009; BYBEE, 2010). As hipóteses a serem consideradas serão avaliadas com evidências experimentais e teóricas. A análise delineada neste estudo trará à tona algumas questões relevantes para o conhecimento sobre a sonoridade da linguagem humana, entre as quais: a interação entre ritmo e cognição, a relação entre padrões silábicos e a atribuição de acento, a consolidação de metodologias de pesquisa em ciências da fala, a fundamentação epistemológica de modelos linguísticos e a imperatividade do trabalho interdisciplinar na formulação de abordagens teóricas.

No segundo capítulo, “Sobre o estudo científico da linguagem”, elabora-se um panorama sobre o estudo da linguagem em uma perspectiva cognitiva, tendo como marco o surgimento da Linguística Cognitiva. As teorias apresentadas são divididas em dois grandes grupos: de um lado, os modelos formais, que veem a linguagem como um sistema computacional autônomo e inato, composto por processos formais que se aplicam serialmente sobre um léxico mínimo. De outro lado, os modelos multirrepresentacionais, que interpretam a linguagem como resultado da interação entre experiência e mecanismos cognitivos gerais, defendendo que as representações linguísticas, detalhadas e redundantes, emergem e se organizam a partir do uso. A comparação entre os modelos formais e os multirrepresentacionais é usada para contextualizar as teorias adotadas neste estudo, bem como para introduzir conceitos teóricos que serão empregados na análise do acento.

O terceiro capítulo, “O acento lexical em abordagens linguísticas tradicionais”, apresenta diferentes análises sobre o acento propostas ao longo do desenvolvimento das teorias linguísticas, fazendo um paralelo com a forma com que estas concebem os mecanismos de funcionamento e as unidades representacionais da dimensão sonora da linguagem. Além disso, nas últimas seções do capítulo, são descritas propriedades funcionais e físicas do acento no português.

O quarto capítulo, “Percepção”, apresenta aspectos básicos da percepção linguística, com enfoque para a complexidade envolvida na percepção da fala, a percepção de aspectos prosódicos e divergências teóricas quanto aos mecanismos perceptuais.

No quinto capítulo, “Abordagem multirrepresentacional do acento”, explora-se a possibilidade de investigar a proeminência acentual no PB com base nos princípios de teorias multirrepresentacionais sobre os sons da fala. Parte-se de uma breve incursão histórica pelo sistema acentual latino, que deu origem ao acento português, para depois propor uma abordagem do acento no português baseada no uso linguístico.

No sexto capítulo, “Análise experimental do acento”, propõe-se uma análise experimental do acento, com a finalidade de dar suporte empírico à análise teórica apresentada nos capítulos anteriores. A primeira parte do capítulo contextualiza a metodologia experimental adotada. A segunda parte é dedicada ao experimento de produção de fala; a terceira, aos três experimentos de percepção. Em cada parte, para cada experimento, são apresentados: critérios metodológicos adotados na coleta de dados, tratamento estatístico dos dados e discussão dos resultados.

Por fim, no sétimo capítulo, “Conclusões”, são retomados os resultados levantados experimentalmente e avaliados diante das reflexões teóricas desenvolvidas nos capítulos iniciais. Propõe-se uma modelagem em redes do fenômeno do acento lexical, que incorpora princípios de funcionamento dos sistemas dinâmicos.

## 2 Sobre o estudo científico da linguagem

Neste capítulo, são discutidos aspectos teóricos do estudo da linguagem humana, tendo como referência a mudança de paradigmas que caracterizou o lançamento de modelos que chamaremos de multirrepresentacionais. São retomadas algumas reflexões apresentadas por Albano (2009), especialmente em sua discussão sobre o percurso de desenvolvimento histórico das teorias linguísticas. Parte-se, inicialmente, do estruturalismo, cujas bases fundadoras foram responsáveis pela consolidação da Linguística como ciência. Em seguida, são apresentados princípios básicos dos modelos formais, tomando como referência o modelo gerativo-transformacional, que entende a linguagem como um sistema computacional autônomo e inato, composto por processos formais que se aplicam serialmente sobre um léxico mínimo. Caracteriza-se, assim, o cenário teórico que serviu de pano de fundo para o surgimento da Linguística Cognitiva e para uma mudança de paradigmas no estudo dos mecanismos cognitivos subjacentes à capacidade linguística. Tal mudança de paradigmas propiciou o desenvolvimento de modelos multirrepresentacionais, que interpretam a linguagem como resultado da interação entre experiência e mecanismos cognitivos gerais, defendendo que as representações linguísticas, detalhadas e redundantes, emergem e se organizam dinamicamente a partir do uso. A comparação entre os modelos formais e os multirrepresentacionais é usada para contextualizar as teorias adotadas neste estudo, bem como para introduzir conceitos teóricos que serão empregados na análise do acento.

### 2.1 Primórdios da linguística

Uma das principais tarefas das teorias linguísticas é a descrição das línguas do mundo em seus diversos aspectos gramaticais. Esta é, sem sombra de dúvidas, uma tarefa árdua, a que têm se dedicado os estudos linguísticos desde seus primórdios, mas que se encontra e sempre se encontrará longe do fim. As línguas mudam ao longo do tempo: a cada dia, a cada século, línguas morrem e línguas nascem. Também os planos espacial e social são cenário de variação e mudança linguística. Uma descrição linguística, portanto, se restringe a determinado grupo social e geográfico, em determinado período de tempo.

As origens da Linguística como ciência dedicada ao estudo da linguagem se confundem com as raízes da epistemologia, da metafísica e da lógica, remontando à Antiguidade Clássica e mesmo a períodos anteriores. Sua consolidação como ciência foi acompanhada, nos séc. XVIII e XIX, pelos estudos filosóficos que refletiam sobre a linguagem como sistema simbólico de expressão e impulsionada pelos estudos filológicos, que realizavam a comparação sistemática entre línguas diferentes para determinar suas relações filogenéticas e estabelecer famílias linguísticas. Inspirando-se nas ciências naturais e no sucesso alcançado pelas teorias de evolução das espécies, os estudos linguísticos preocupavam-se em estabelecer princípios de funcionamento da linguagem – e.g. a formulação de regras sonoras responsáveis pela transformação temporal das línguas, vistas como sistemas em desenvolvimento. As mudanças sonoras, assim como os elementos ou unidades atômicas dos sistemas sonoros, passaram a atrair especial atenção. A partir dos métodos desenvolvidos pelos estudos histórico-comparativos, floresceu a tradição estruturalista, que se aprofundou, não só nos aspectos sonoros, mas em outras dimensões da linguagem (sintaxe, semântica, etc.), e conferiu grande impulso à descrição das línguas do mundo<sup>1</sup>.

Nesse mesmo período, tomou forma a ideia de que, a partir de um número suficiente de descrições linguísticas pormenorizadas, seria possível isolar mecanismos e padrões gerais, “universais linguísticos”, que, ocorrendo sistematicamente em todas as línguas do mundo, poderiam ser interpretados como características da linguagem humana em geral. Impôs-se a tarefa de identificação e elaboração de conceitos, com a finalidade de homogeneização de metodologias e terminologias. Surgiram perguntas como: quais são as unidades conceituais básicas da linguagem humana? Quais as características acústicas e articulatórias dos sons da fala?

As teorias científicas buscam, não só a descrição dos fatos, mas também, quando possível, uma explicação acurada da realidade em termos de leis universais, de um sistema teórico coerente, contemplando as relações causais de eventos e estruturas. Novas perguntas surgiram: como se formam as unidades linguísticas? Como se organizam no plano físico da fala? Como se organizam na mente dos falantes? Não bastava à tarefa linguística simplesmente nomear um evento ou estrutura; tornou-se evidente a necessidade de compreender os fatores e mecanismos que subjazem aos mesmos. Acompanhando o

---

<sup>1</sup> Para um panorama histórico da linguística, cf. Bynon (1986), Geeraerts (2010) e Parret (1975).

estabelecimento do método científico (cf. POPPER, 1968; SCHLICK et al., 1975), difundiu-se uma preocupação em construir teorias bem fundamentadas, testáveis (i.e. refutáveis) e com poder explanativo, que conseguissem modelar os fenômenos linguísticos de forma sistemática e abrangente, estabelecendo previsões e generalizações.

## 2.2 Modelagem formal: a linguagem como sistema serialista computacional

A busca pelos universais linguísticos sempre foi tema de grande fascínio para os linguistas por todo o mundo, especialmente porque chega a representar, na linguística, uma busca pelo “elo perdido”: os elementos universais da linguagem seriam um dos caminhos para a elucidação das circunstâncias em que teria surgido a capacidade linguística exclusiva e generalizada na espécie humana. O estudo dos universais linguísticos, no entanto, é objeto de grandes controvérsias e debates. Se, de um lado, conseguiu-se construir um corpo respeitável de descrições detalhadas de diversas línguas do mundo (e.g. GREENBERG, 2005a, 2005b; LADEFOGED; MADDIESON, 1996), de outro, o estabelecimento de regularidades ou tipologias só se mostrou possível em termos de tendências com maior ou menor grau de generalidade, e não de premissas sem exceções.

Levando ao extremo a ideia de elementos universais, foi traçada uma via de análise com objetivo de modelagem de um sistema linguístico único, inato, composto por princípios básicos de uma gramática latente, responsável por gerar as estruturas da linguagem atualizada, no plano físico (e.g. CHOMSKY, 1972; PINKER, 1994). Tal vertente teórica, em repúdio aberto a teorias behavioristas do aprendizado e recrudescendo antigo embate epistemológico, adota um posicionamento racionalista que culminou em uma intrincada e sofisticada abordagem matemático-computacional da linguagem. Por meio do raciocínio lógico inferencial de base intuitiva, o conhecimento linguístico implícito, idealizado, foi concebido sob a forma de um sistema simbólico, composto por módulos e regras. A teoria gerativo-transformacional, uma das mais renomadas nesse paradigma de pesquisa, busca demonstrar como a linguagem humana poderia ser gerada a partir de uma gramática formal de tipo transformacional (CHOMSKY, 1956), constituída de regras e provida de um léxico mínimo.

A teoria gerativo-transformacional teve o mérito de demonstrar que a linguagem não pode ser resultado exclusivamente da interação entre o homem e o ambiente, como pretendia a teoria behaviorista ao propor que o comportamento humano (inclusive o linguístico) seria redutível a um sistema estímulo-resposta. É necessário haver algum tipo de substrato cognitivo e mecanismos de processamento ativo da linguagem, que expliquem por que a estrutura linguística é criativa, adquirida por todo ser humano em condições normais de desenvolvimento e apenas pela espécie humana. Tais argumentos foram utilizados para construir a teoria nativista do desenvolvimento linguístico, especialmente advogada pela teoria gerativa, cognitivista, e por esta atrelada a um modelo baseado em regras, algoritmos de manipulação simbólica e cognição interna. A linguagem, tendo como núcleo a sintaxe, é vista como inata e dotada de um módulo cognitivo autônomo (cf. FODOR, 1995; PINKER, 1994). Como veremos adiante, contudo, ao tratar dos modelos multirrepresentacionais, há outras possibilidades de explicação dos argumentos contra o behaviorismo, mas que dispensam o recurso a uma capacidade linguística inata subjacente, ao se considerar que a estrutura linguística emerge a partir de processos cognitivos gerais.

Ainda que a modelagem da linguagem humana por meio de gramáticas formais tenha alcançado grande sucesso na representação da estrutura sintática observada em várias línguas do mundo, ela mesma reconhece sua fragilidade em modelar outros aspectos da linguagem humana, como o sonoro e, principalmente, o semântico. Além disso, o fato de que, alegadamente, o modelo gerativo-transformacional consegue gerar com sucesso as estruturas sintáticas de uma língua não prova, por si só, que este é o mecanismo cognitivo de fato responsável pelas referidas estruturas. Como veremos à frente, outras possibilidades teóricas se abrem à modelagem da linguagem humana, algumas das quais mais robustas, pois se submetem melhor aos princípios de réplica e refutabilidade

Chomsky (1995) reconhece que, de um lado, a linguagem humana, em seu uso, é um sistema biológico e, como tal, apresenta estrutura caótica, intrincada e moldada por circunstâncias acidentais; em tais sistemas, a redundância, esperada, seria responsável por compensar danos e imperfeições e por acomodar fins e funções variados. Por outro lado, a despeito desse quadro, o autor argumenta que, “in its basic structure, the language faculty has properties of simplicity and elegance that are not characteristic of complex organic systems, just as its infinite digital character seems biologically rather isolated.” (CHOMSKY, 1995, p. 29), ainda que assumindo, em seguida, a possibilidade de que a interpretação computacional

possa ser um artefato refletindo seus próprios métodos de pesquisa. Restringir a teoria da faculdade linguística à competência de um falante ideal é prejudicial à própria ideia de universalidade pretendida pela teoria gerativa.

A partir principalmente da década de 1960, delineou-se uma mudança de paradigma. Os estudos sobre a linguagem humana se voltaram para a construção de uma teoria linguística que explicasse o funcionamento e a interação de cada dimensão da linguagem, bem como sua implementação biológica e social, de forma integrada com outros processos mentais (cf. ELLIS; LARSEN-FREEMAN, 2009; GELDER; PORT, 1995; LANGACKER, 1987, entre muitos outros). Tal empreendimento, contudo, mostrou necessário reunir conhecimentos e descobertas de múltiplos campos de pesquisa, como a psicologia, sociologia, física, neurociências e tecnologia computacional, a fim de construir uma teoria linguística em conformidade com os fatos reais da linguagem humana. Impulsionada por avanços tecnológicos e científicos, provenientes de várias áreas – como computação, estudos de *corpora*, modelagem matemático-computacional em redes, inteligência artificial, genética, neurofisiologia, motricidade, comportamento humano e animal –, formou-se o que se pode chamar de segunda geração da ciência cognitiva, propondo teorias baseadas em esquemas, probabilidade e cognição social (cf. SINHA, 1988, 2007). Batizada como Linguística Cognitiva, essa nova perspectiva teórica diverge ortogonalmente da primeira geração, a abordagem cognitivista clássica representada pelo gerativismo, com a qual compartilha não mais que o interesse pela esfera cognitiva dos eventos e estruturas linguísticos. Na seção a seguir, buscaremos reconstruir o contexto teórico de surgimento da Linguística Cognitiva, tendo como eixo central o questionamento a dois pressupostos fundamentais da teoria gerativo-transformacional: o modularismo e o inatismo.

### 2.3 Surgimento da Linguística Cognitiva

Cognitive linguistics is the study of language in its cognitive function, where *cognitive* refers to the crucial role of intermediate informational structures in our encounters with the world. Cognitive Linguistics is cognitive [...] by assuming that our interaction with the world is mediated through informational structures in the mind [...] focusing on natural language as a means for organizing, processing, and conveying that information. Language, then, is seen as a repository of world knowledge, a structured collection of

meaningful categories that help us deal with new experiences and store information about old ones. (GEERAERTS; CUYCKENS, 2007, p. 5) .

Do ponto de vista epistemológico, o surgimento da Linguística Cognitiva não foi simplesmente uma reação ao racionalismo extremado da teoria gerativo-transformacional, mas também deu continuidade a pesquisas em psicologia desenvolvidas décadas antes. Como descreve Matlin (1994), o behaviorismo predominou nos Estados Unidos como corrente teórica em psicologia durante as décadas de 1930 e 1940, voltando-se para a primazia dos estímulos ambientais e externos, restringindo-se à análise do comportamento observável, interessada em mecanismos de condicionamento. Enquanto isso, entretanto, desenvolvia-se na Europa uma perspectiva de análise fundamentalmente distinta: a psicologia da Gestalt direcionou sua atenção para tendências humanas de organização, propondo uma organização inerente ao todo na experiência, dando ênfase ao estudo da percepção de padrões, principalmente visuais, que originaram conceitos consagrados como figura-fundo e perspectiva. A Linguística Cognitiva, em sua formulação recente, retoma e desenvolve as ideias da psicologia Gestalt, como se pode verificar, por exemplo, na sistematização de conceitos realizada por Croft e Cruse (2004). Tais conceitos se mostram válidos para explicação e interação de vários processos cognitivos, entre os quais a linguagem, e recebem suporte de estudos recentes sobre a estruturação e o funcionamento do cérebro, como veremos adiante.

Ainda que as teorias inatistas e modularistas da linguagem tenham alcançado grande sucesso e muitos adeptos na comunidade linguística, argumentos e evidências podem ser apontados a favor de uma visão da linguagem humana como fenômeno emergente da interação entre processos cognitivos gerais e o uso linguístico.

Uma das críticas dirigidas especificamente ao paradigma gerativo diz respeito à excessiva centralidade conferida aos processos de hierarquização e recursividade em face da escassez de propriedades linguísticas específicas sistematicamente compartilhadas pelas línguas do mundo e da atestada fragilidade dos “universais” linguísticos.

If someone asked ‘What properties do rose bushes share that enable us to recognize them as rose bushes?’, it would be ridiculous to answer ‘They are hierarchically structured.’ If knowledge of language were innate in humans as a product of biological evolution, human languages ought to share a mass of common properties so specific and concrete that someone discussing language universals would never think of mentioning hierarchical structure. (SAMPSON, 2005, p. 150).

Nos estudos culturais e antropológicos, o questionamento do modularismo e do inatismo linguísticos, em geral, se vale do fato de que (1) não se tem conhecimento, até os dias atuais, de nenhuma estrutura específica para a linguagem em todo o corpo humano, do código genético à arquitetura cerebral, passando pela anatomia do aparelho fonador, e (2) o surgimento da modularidade linguística na mente humana não pode ser claramente delineado em termos evolutivos. O primeiro argumento é explicado pelo fato de que a linguagem coopta processos cognitivos gerais, como memória, raciocínio, atenção, categorização, emoções, que têm lugar no sistema nervoso central. Coopta também outros sistemas corporais, envolvidos diretamente na expressão sonora e gestual, como sistema respiratório, auditivo, face, mãos, além de afetar indiretamente todo o corpo, que entra em sintonia com ritmo e unidades da fala. Conectando o primeiro ao segundo argumento contra o modularismo e inatismo, Tomasello (1999, p. 2) defende que os seis milhões de anos que separam os homens dos outros primatas superiores são pouco tempo em termos evolutivos para justificar o surgimento de estruturas biológicas específicas:

The fact is, there simply has not been enough time for normal processes of biological evolution involving genetic variation and natural selection to have created, one by one, each of the cognitive skills necessary for modern humans to invent and maintain complex tool-use industries and technologies, complex forms of symbolic communication and representation, and complex social organizations and institutions.

Na filogênese humana, as línguas teriam evoluído em complexidade ao longo do tempo, paralelamente ao desenvolvimento cultural e social de seus povos. Beckner et al. (2009, p. 14), remetendo a estudos recentes, concluem que

The dynamics in the population of linguistic utterances is therefore not connected in a simple way to that of the underlying human population; hence, the role of natural selection in shaping languages is likely to be diminished compared to what has sometimes been assumed elsewhere.

Somente pela transmissão cultural ou social seria possível explicar mudanças cognitivas e comportamentais tão rápidas (cf. TOMASELLO, 1999 para uma hipótese sobre o surgimento, nos eixos ontogenético e filogenético, da capacidade linguística por meio da transmissão cultural<sup>2</sup>; cf. DUNN et al. 2011, que, em uma avaliação tipológica sobre ordem

<sup>2</sup> Tomasello (1999) atribui a singularidade da cognição humana ao desenvolvimento, no plano filogenético, da capacidade de autoidentificação com coespecíficos, o que, no plano histórico, teria favorecido novas formas de aprendizagem cultural e acúmulo de modificações em artefatos e tradições, em meio aos quais nasce e se desenvolve a criança humana. Esta, por sua vez, tem à sua disposição, eixo ontogenético, um conjunto

das palavras, concluem pela transmissão cultural da linguagem; cf. STEELS, 2011 para uma aplicação da inteligência artificial na modelagem dos mecanismos de transmissão cultural e genética da linguagem).

Já foram descritos em detalhes diversos sistemas não humanos de comunicação e cultura, alguns com elevado nível de complexidade, o que, segundo Sinha (2005), apontaria para (mas não implicaria) uma possível gradualidade na evolução da capacidade linguística na espécie humana. Em vista dessa consideração, o autor, posicionando-se no intermédio entre empirismo e racionalismo, concebe a capacidade da linguagem na forma de emergência epigenética e elaboração simbólica.

[...] in contrast to non-human signal systems of communication, human natural languages are symbol systems. The evolutionary transition from signal to symbol usage, and the exosomatic, culturally-driven elaboration of symbol usage into language, accounts for the unique complexity of human language (including grammar). This emergent complexity, I suggest, has, in the course of evolution co-opted or captured a suite of cognitive capacities that are uniquely developed (but not unique) in humans. There is no contradiction, I am claiming, between recognizing both the qualitative uniqueness of human language, and the essential continuity between human and non-human neurobiology. (SINHA, 2005, p. 313).

Em vista dessas considerações, a existência de universais linguísticos pode ser aceita se relativizada, ou seja, se os mesmos forem interpretados como tendências gerais que emergiram em consequência de condições singulares compartilhadas pela espécie: condições que são tanto biológicas (características fisiológicas, anatômicas, cognitivas), quanto sociais (transmissão cultural do conhecimento, estrutura de organização social) e ecológicas (*habitat* na superfície da Terra). Consequentemente, explica-se como a linguagem – tenha ela ou não uma origem única – se desenvolveu em todos os povos usando os mesmos órgãos e a partir da mesma necessidade comunicativa, envolvendo as mesmas ações, em mesmo ambiente físico (PALMER, 1999). As diferenças entre as línguas decorreriam em grande parte da variação individual e cultural. Não segue dos “universais”, portanto, a conclusão de que a linguagem humana resulta de um módulo inato específico da espécie, se é possível atribuí-los ao compartilhamento de um conjunto extraordinário formado pelas estruturas cognitivas de alto nível e pelos sistemas corporais adequados à linguagem.

---

compartilhado de conhecimentos e habilidades das quais pode se beneficiar e que lhe facultariam tanto a aquisição e uso de representações cognitivas perspectivas para formar símbolos linguísticos, quanto a internalização de interações discursivas.

As yet, we know too little to accept or reject hypotheses regarding the innateness of a specifically syntactic component of the human language faculty. I certainly would not wish to reject the possibility that the epigenetic processes selected for in human evolution include a predisposition for learning language, although this does not necessarily imply that any such predisposition is or was “dedicated” from the start exclusively to language. (SINHA, 2005, p. 333).

Do ponto de vista filogenético, portanto, não há evidências conclusivas nem que comprovem nem que rejeitem a hipótese de que a linguagem humana resulta de um módulo autônomo inato e exclusivo à espécie, mas muitos são os estudos que argumentam pelo papel decisivo e inegável da transmissão cultural do conhecimento no surgimento e desenvolvimento da linguagem humana. Contudo, pode-se atribuir a falta de evidências que esclareçam o advento da capacidade linguística à natureza indireta das investigações sobre a filogênese da espécie, sendo necessário, portanto, buscar respaldo em fontes de outra natureza, como pesquisas no eixo ontogenético e em neurociências.

Na psicologia do desenvolvimento, também não há consenso quanto à natureza da capacidade da linguagem, ainda que sejam inúmeras as evidências que tendem ao não modularismo e inatismo. É fato conhecido que “[...] for language acquisition to be successful, the child’s linguistic input must be in the form of interactive, social speech.” (CAIRNS, 1986, p. 6). Sabe-se que a aquisição da linguagem acontece simultaneamente e inter-relacionadamente com a aquisição de outros processos cognitivos, principalmente envolvidos na capacidade simbólica e de conceitualização. Tais processos acontecem necessariamente ao longo da interação da criança com o meio cultural e social. Vigotski (2001), sumariando observações e pesquisas experimentais com animais e crianças, aponta que a relação entre o pensamento e a linguagem, tanto no eixo filogenético quanto no ontogenético, é uma grandeza variável, que oscila ao longo do desenvolvimento e se dá de forma não paralela e não homogênea. Pensamento e linguagem têm raízes diferentes e, até certo ponto de seu desenvolvimento, as duas modalidades apresentam trajetórias distintas e independentes. Tanto nos antropóides quanto nas crianças em fase inicial de desenvolvimento (cerca de um ano), pode-se identificar um estágio pré-verbal do pensamento e um estágio pré-intelectual da linguagem, em que esta tem função, não só emocional, mas também social. Nas crianças, contudo, observa-se, por volta dos dois anos de idade, uma mudança que não apresenta paralelo nos antropóides: convergem as curvas de evolução do pensamento e da linguagem, que até então tinham trajetórias independentes. Neste momento da ontogênese humana, em

que é sintomático o aumento ativo do vocabulário de forma rápida e aos saltos, “a fala se torna intelectual e o pensamento verbalizado” (VIGOTSKI, 2001, p. 131). Por meio da experiência sociocultural, há, nesse estágio do desenvolvimento verbal, uma assimilação puramente externa do signo (VIGOTSKI, 2001, p. 146), em que a palavra é interpretada pela criança como uma propriedade do objeto nomeado, estágio necessário para que, anos depois, ela seja capaz de descobrir a função simbólica da linguagem. Conclui que “o pensamento verbal não é uma forma natural e inata de comportamento, mas uma forma histórico-cultural” (VIGOTSKI, 2001, p. 149).

Sinha (2006, p. 166 e 168) dá ainda maior ênfase à interação entre a criança e o meio social no desenvolvimento das habilidades linguísticas:

[...] linguistic meaning [is] socially and discursively constructed on the basis of both subjective, schematized experience, and intersubjective, shared reference. [...] the development of language and cognition is not only *schematically, iconically, and non-discursively grounded* in embodied experience; it is *also discursively grounded* in an intersubjectively constructed and socially-shared, signified world.

De fato, o conhecimento de uma palavra não consiste na sua simples associação com um número específico de significados estáveis, mas antes no domínio ativo da extensão altamente flexível de uso desta palavra (TAYLOR, 2006). Estudos apontam para a dinamicidade das unidades linguísticas (e.g. BYBEE, 2010) e sua organização em redes de probabilidades (HAY et al., 2004; PIERREHUMBERT, 2003). Beckner et al. (2009, p. 5-6) mostram que “both infants and adults track co-occurrence patterns and statistical regularities in artificial grammars. Such studies indicate that subjects learn patterns even when the utterance corresponds to no meaning or communicative intentions”. Sabe-se que, como os humanos, alguns animais dispõem de uma noção probabilística acurada (SILVERMAN, 2006). Portanto, a organização probabilística do conhecimento linguístico seria mais uma propriedade da linguagem humana não específica da espécie.

Passando às pesquisas em neurociências, a teoria localizacionista em sentido estrito veio, nas últimas décadas, perdendo cada vez mais espaço frente à compreensão crescente do funcionamento cerebral, sua plasticidade e supercapacidade de processamento simultâneo de informações (ELMAN et al., 1996).

Research over the past decades has clearly established that neural circuits in the brain undergo continuous, dynamic modification in relation to learning

and memory formation, in the form of synaptic strengthening and weakening, synapse formation and remodeling, and neurogenesis. (BRUEL-JUNGERMAN et al., 2007)

Além disso, sabe-se hoje que, na estrutura e fisiologia cerebrais, há enorme redundância quanto ao número de ligações sinápticas entre neurônios. Por que não haveria redundância também no armazenamento de informações, na representação dos eventos linguísticos e de quaisquer outros eventos processados cognitivamente?

Em geral, argumentos pelo localizacionismo estrito e modularidade se valem de estudos com pacientes com lesão cerebral buscando identificar uma relação entre a região lesionada e as funções comprometidas. Esse tipo de estudo, contudo, se relega à descrição de casos individuais, sem o poder de respaldar afirmações gerais, universais, uma vez que dados de patologias de fala decorrentes de comprometimento cerebral podem ser reflexo de uma reorganização cerebral pós-lesão, do desenvolvimento de estratégias de compensação para suprir as funções comprometidas, ou mesmo de variações individuais estabelecidas previamente à ocorrência do dano (BATES et al., 1991). Ademais, diversas são as pesquisas em neurofisiologia que questionam a possibilidade de se estabelecer uma correspondência simples entre tarefas complexas e áreas cerebrais específicas.

Pesquisas com primatas indicam que a reorganização neural, mesmo em adultos, pode se estender por centímetros na área somatossensorial do córtex, passados alguns anos da amputação de um membro, abrindo a possibilidade de duas explicações: “New wiring can be manufactured and established in the adult brain, or the old patterns of connectivity can be converted [...] to serve functions that they never served before.” (ELMAN et al., 1996, p. 3). Com isso, a hipótese localizacionista perde força como argumento em prol do inatismo.

Os neurônios-espelho, tipo de neurônio ativado tanto ao realizar quanto ao observar uma ação, teriam papel fundamental na estruturação da linguagem, ao participarem da codificação neural ou corporificação da experiência com o mundo (ARBIB, 2010).

Current information about the structure and development of the human brain offers no easy solutions to the nature-nurture problem. [...] It's interaction all the way down, from top to bottom, and apparent universals often mask a range of alternative outcomes that can be demonstrated when default input conditions do not hold. (ELMAN et al., 1996, p. 251).

Em resumo, foi-se reunindo, especialmente nas últimas três décadas, um conjunto variado de dados e evidências que levam a uma concepção de linguagem de natureza muito

diversa da que era delineada pelos modelos formalistas computacionais: como vimos acima, pesquisas em filogênese, ontogênese, funcionamento e organização cerebrais não respaldam a concepção modularista-inatista da linguagem humana. Nas três áreas de estudo, há indícios reiterados que apontam, antes, para a emergência da linguagem a partir da interação de estruturas e mecanismos inatos (mas não específicos) com a experiência e o ambiente.

Esforços vêm sendo feitos com intuito de sustentar e atualizar o paradigma gerativista em face de descobertas recentes sobre genética e psicologia experimental como as apresentadas acima. Seria uma tentativa de conferir plausibilidade orgânica à teoria, em uma espécie de corporificação de um modelo que, a princípio, era completamente desvinculado das estruturas corporais e cerebrais. Em um artigo relativamente recente (HAUSER; CHOMSKY; FITCH, 2002), a faculdade da linguagem é definida como composta de um sistema sensório-motor e um sistema conceitual-intencional acessórios a um sistema nuclear, de natureza computacional e abstrata, independente de outros sistemas, e cuja estrutura interna contém um sistema computacional (sintaxe estrita, grande responsável pela recursividade, infinitude discreta), um sistema fonológico (mapeamento sensório-motor) e um sistema semântico (mapeamento conceitual-intencional). Essa pode ser vista como uma tentativa de abertura teórica a descobertas recentes na área de neurociências e uma reação ao sucesso alcançado pelos modelos da segunda geração da Linguística Cognitiva. Ou se restringe a faculdade da linguagem como módulo inato exclusivamente à recursividade ou permanecem os problemas anteriores relacionados à natureza serialista, computacional e modular da teoria, cuja superação levaria mesmo a uma ruptura com os pressupostos nucleares do gerativismo e a uma completa reformulação teórica.

O gerativismo passou a dividir terreno com teorias que, como a Linguística Cognitiva, se destacaram nas últimas décadas por ver a linguagem como resultado de processos cognitivos gerais em interação com a experiência e por tentar encontrar um ponto de equilíbrio frente ao debate *nature x nurture*. Tais teorias podem ser reunidas em um conjunto, sob o nome de modelos multirrepresentacionais (cf. CRISTÓFARO-SILVA; GOMES, 2007). Entre estes modelos, são de interesse do presente estudo, além da Linguística Cognitiva, os Modelos de Uso e os Sistemas Dinâmicos. Há evidências de que a integração de tais modelos teóricos não só é viável, como proveitosa, visto que se suplementam: Modelos de Uso e Linguística Cognitiva serão aqui empregados na descrição e explicação de fenômenos linguísticos, sendo amparados pelos Sistemas Dinâmicos.

## 2.4 Abordagens multirrepresentacionais

Destacaram-se nas últimas décadas teorias linguísticas que assumem que as representações linguísticas são múltiplas e que a linguagem emerge a partir do uso. Tais teorias compartilham uma série de pressupostos que divergem das teorias linguísticas tradicionais e podem ser agrupadas como modelos multirrepresentacionais. A FIG. 1, a seguir, sistematiza as principais diferenças entre modelos tradicionais e multirrepresentacionais na análise dos sistemas sonoros, que serão de especial interesse neste estudo.

<b>Proposta tradicional</b>	<b>Modelos multirepresentacionais</b>
Representação mental minimalista	Representação mental detalhada
Separação entre fonética e fonologia	Inter-relação entre fonética e fonologia
Visão da fonologia como uma gramática formal, com a utilização de variáveis abstratas	Consideração de que a fonologia da língua envolve a distribuição probabilística de variáveis.
Efeitos da frequência refletidos na produção em curso e não armazenados da memória de longo termo.	Efeitos da frequência armazenados na memória de longo termo.
Julgamento fonotático categórico: uma seqüência ou é considerada bem formada ou é impossível de ocorrer na língua.	Efeitos gradientes nos julgamentos fonotáticos
Léxico separado da gramática fonológica	Palavra como locus da categorização

FIGURA 1 - Comparação entre modelos linguísticos tradicionais e multirrepresentacionais

FONTE: OLIVEIRA, 2003 apud CRISTÓFARO-SILVA; GOMES, 2007, p. 164.

Os princípios dos modelos multirrepresentacionais sistematizados na FIG. 1 serão de fundamental importância para a análise do acento lexical aqui proposta e serão retomados ao longo dos capítulos desta tese. O presente estudo irá realizar uma análise do acento lexical com base em modelos multirrepresentacionais, mais especificamente os Modelos de Uso, a Linguística Cognitiva e Sistemas Dinâmicos. A seguir passaremos a descrever os princípios básicos de cada um.

A modelagem da linguagem como um sistema dinâmico é uma perspectiva que tem se fortalecido nas últimas décadas, principalmente diante da abertura da linguística para o trabalho interdisciplinar. A teoria dos Sistemas Dinâmicos foi desenvolvida na matemática

para explicar o comportamento de conjuntos de componentes em mudança ao longo do tempo e que se influenciam mutuamente (GELDER; PORT, 1995). A modelagem dinâmica deste tipo de comportamento tem como característica central o tempo tratado como variável do sistema. O estado atual ou final de um sistema dinâmico envolve a determinação de um estado inicial, sobre o qual operam forças do sistema e atratores (estados a que tendem a convergir as variáveis do sistema). Sistemas dinâmicos, portanto, apresentam grande sensibilidade aos estados iniciais e suas variáveis tendem a se estabilizar temporariamente em torno de atratores. Fenômenos que apresentam comportamento dinâmico são formados sistemas embutidos (*embedded*) e interconectados, de forma que mudanças em uma variável do sistema exerce influência nas outras variáveis, o que é determinante para a auto-organização do sistema. Uma das preocupações centrais da Teoria dos Sistemas Dinâmicos diz respeito aos padrões de coordenação – como os estados de coordenação de um sistema, observados por meio de variável relevante, evoluem ao longo do tempo – e como se desenrolam as interações sutis e complexas entre partes do sistema (KELSO, 1995).

Diversos aspectos do comportamento humano podem ser bem descritos por meio de Sistemas Dinâmicos, desde a motricidade até processos cognitivos, passando pela linguagem (cf. GELDER; PORT, 1995; THELEN; SMITH, 1994). Os Sistemas Dinâmicos têm muito a contribuir com uma abordagem multirrepresentacional da linguagem humana, especialmente por oferecerem um tratamento não categórico à complexidade nos sistemas linguísticos, vista como emergente a partir da repetição de eventos, muitas vezes envolvendo pequenas mudanças em variáveis do sistema. Outro aspecto fundamental da abordagem dinâmica para análise linguística diz respeito ao papel do tempo, que transcorre de forma imediata, sendo possível considerar a evolução real do sistema linguístico, por padrões de mudança não linear, no eixo temporal, não como uma mera sucessão de estados estáticos.

Uma das perspectivas dinâmicas da fala modela matematicamente a coordenação dos articuladores no trato vocal em sua movimentação que se desdobra temporalmente e espacialmente (cf. FUCHS et al., 2012). Outro exemplo de abordagem dinâmica em linguística, de especial interesse para o presente estudo, argumenta que a linguagem pode ser vista como um sistema adaptativo complexo (cf. BYBEE, 2010; BECKNER et al., 2009 e demais trabalhos em ELLIS; LARSEN-FREEMAN, 2009). Neste trabalho, serão adotados princípios fundamentais dos Sistemas Dinâmicos em suporte à abordagem baseada nos Modelos de Uso e na Linguística Cognitiva.

Os chamados Modelos Baseados no Uso (*usage-based models*), ou simplesmente Modelos de Uso, compartilham a ideia de que os sistemas linguísticos são construídos a partir de eventos de interação linguística e uso da linguagem, sendo que a relação entre estas duas dimensões – sistema e uso – é relativamente direta (BARLOW; KEMMER, 2000). O termo “*usage-based*”, que remonta a Langacker (1987) e é explorado em Langacker (2000), refere-se a uma abordagem da linguagem que se contrapõe à perspectiva formalista tradicional por ser maximalista (assumindo que a representação cognitiva da linguagem envolve grande redundância e detalhamento), não redutiva (incorporando à gramática tanto estruturas gerais e abstratas quanto unidades linguísticas particulares) e *bottom up* (defendendo que as estruturas gerais abstratas não existem *a priori*, mas emergem na forma de esquematizações a partir de instâncias de uso linguístico).

Na Linguística Cognitiva, que é um modelo baseado no uso, a estrutura simbólica é fundamentada (*grounded*) em eventos de uso e em rotinas motoras de vocalização, como mostra a FIG. 2.

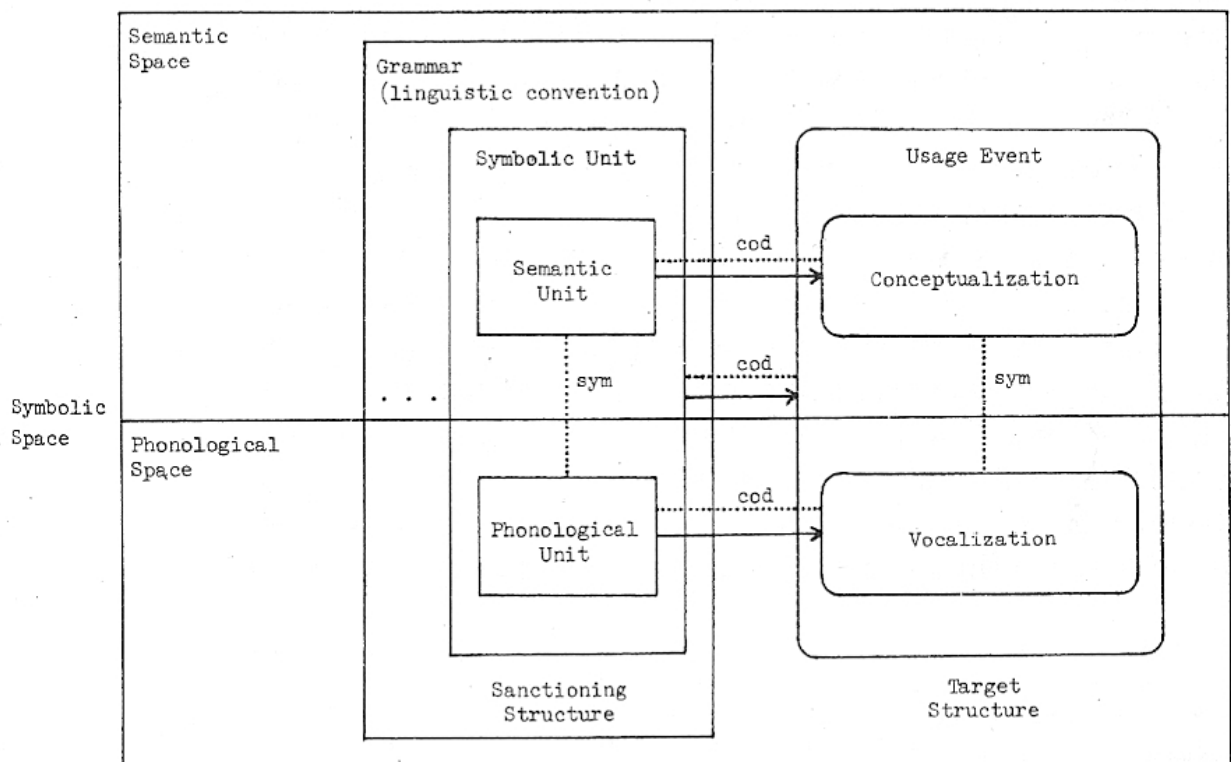


FIGURA 2 - Estrutura simbólica da linguagem segundo a Linguística Cognitiva

FONTE: LANGACKER, 1987, p. 77.

Em resumo, as abordagens de uso, como a Linguística Cognitiva, assumem que, no armazenamento de informações, um grande volume de dados estaria envolvido – por isso seriam maximalistas, e não minimalistas. Nesse armazenamento, haveria grande redundância de informações – razão pela qual esses modelos são chamados de não redutivos. Desse modo, na mente humana, a especificidade estaria atrelada à generalidade, sendo que a última surge a partir da primeira – daí se dizer que Modelos de Uso envolvem primordialmente um processamento *bottom up*. Formulações muito próximas a essas têm sido adotadas, principalmente a partir da década de 1980, por estudos tanto independentes quanto derivados da Linguística Cognitiva, mas todos eles contribuindo para fortalecer as bases gerais da abordagem de uso. Especialmente nas duas últimas décadas, vem-se acumulando um conjunto respeitável de evidências que indicam a primazia do uso em áreas como psicologia, linguística, cognição e neurofisiologia. Nessas áreas, destacam-se estudos pioneiros sobre processos cognitivos e de categorização, comprometidos com a linguagem humana em graus diferentes, mas em grande medida aplicáveis a seu estudo e compatíveis entre si.

Como vimos, os Modelos de Uso e a Linguística Cognitiva defendem que a linguagem emerge e é moldada pelo uso. Outro pressuposto fundamental da Linguística Cognitiva compartilhado pelos Modelos de Uso é a visão da linguagem como produto de vários processos cognitivos que têm lugar na mente humana, com os quais está diretamente relacionada – ao contrário das teorias modularistas da linguagem, que defendem a existência de um módulo linguístico específico e autônomo responsável pela faculdade humana da linguagem. Busca-se, antes, explicar a linguagem integrando-a a outros mecanismos característicos da mente humana, como armazenamento e processamento de informações, categorização, iconicidade (cf. GEERAERTS, 2006).

O presente estudo se fundamentará em duas abordagens de uso relativamente recentes, voltadas para a dimensão sonora da linguagem: a Fonologia de Uso (BYBEE, 2001, 2010) e o modelo de exemplares proposto em Pierrehumbert (2001a). Estas duas abordagens são complementares à proposta de Langacker (1987), outro ponto de referência teórica neste estudo que, ao lado de Lakoff (1987), foi um dos grandes responsáveis pela repercussão alcançada pela Linguística Cognitiva nas últimas décadas. Alguns conceitos fundamentais compartilhados pela Linguística Cognitiva, Modelos de Uso e de Exemplares serão brevemente sumariados a seguir, com base nas propostas de Bybee (1985, 1995, 2001, 2010), Langacker (1987), Phillips (2001, 2006), Pierrehumbert (2001a, 2003).

A ideia de que as representações mentais são estruturadas em formas de exemplares foi primeiro proposta em estudos psicológicos, como Medin e Schaffer (1978, teoria contextual de classificação baseada na recuperação de informações dos exemplares armazenados), posteriormente desenvolvida por Nosofsky (1986, modelo quantitativo de identificação e categorização de estímulos perceptuais multirrepresentacionais) e Hintzman (1986, modelo de memória episódica MINERVA 2 baseado na abstração de esquemas a partir do armazenamento de traços múltiplos e explorado em uma perspectiva linguística por Goldinger (1997)). Modelos de exemplares que, como os mencionados acima, simulam fenômenos de percepção e categorização estão na base de modelos linguísticos como, por exemplo, o de Johnson (1997) e o de Pierrehumbert (2001a, modelo de exemplares para percepção e produção com base em uma gramática estocástica). Os modelos de exemplares guardam algumas semelhanças com o modelo de protótipos, de Rosch (1978), na psicologia, e com as categorias radiais, de Lakoff (1987), na linguística. Esses três modelos representacionais podem ser considerados desenvolvimentos análogos na investigação das categorias mentais humanas, opostos ao modelo aristotélico clássico, que parte de unidades discretas e bem delimitadas (TAYLOR, 1991). Protótipos diferem dos exemplares, contudo, quanto à estrutura interna: são gradientes, centrados em membros nucleares bem definidos e difusos nas bordas. Já os exemplares não pressupõem a existência de hierarquização interna e podem sofrer sobreposição entre categorias. Exemplares são representados por nuvens de memórias cognitivas detalhadas, linguísticas e não linguísticas, organizadas em uma rede de relações de naturezas variadas. Tais relações possuem uma determinada força de conexão associada (*lexical strength*).

Os estudos voltados para a dimensão sonora mostram que o detalhe fonético fino é armazenado e desempenha uma função crucial tanto no processamento quanto na armazenagem das informações (HAWKINS, 2003; HAY et al. 2004). A categorização acontece por meio das conexões entre exemplares, a partir de relações de semelhança e extensão, e seu resultado não são categorias discretas, mas um contínuo: essa abordagem consegue explicar fenômenos de gradualidade envolvidos na categorização linguística. Por outro lado, o espaço perceptual exhibe efeitos de granularidade (e.g. mínima diferença percebida; percepção categórica), que podem ser modelados em função da acuidade ou resolução do sistema perceptivo – uma vez que a linguagem é corporificada, valendo-se de mecanismos cognitivos gerais e sujeitas a restrições fisiológicas e funcionais do organismo.

Diversas teorias linguísticas atuais concordam que o que tradicionalmente se concebeu como “léxico” pode ser interpretado como uma rede de ligações cognitivas entre exemplares, compostos por informações de diversas naturezas, entre as quais: *construals*<sup>3</sup>, imagens, sons, emoções, palavras, *chunks*<sup>4</sup> linguísticos. Bybee (2002a, p. 112) aponta quatro tipos de evidência de que as memórias linguísticas envolveriam a armazenagem de *chunks* desencadeada a partir da alta frequência com que determinadas sequências ocorrem:

First, one unit of a chunk primes or automates the other unit. [...] Second, inside frequently used chunks, internal structure tends to be lost. [...] Third, the morphemes or words inside a chunk become autonomous from other instances. [...] And fourth, the components of a chunk become compressed and reduced phonologically.

A autora generaliza tais processos com a hipótese de que “items that are used together fuse together” (BYBEE, 2002a, p. 112), plenamente compatível com a teoria de funcionamento neuronal no aprendizado associativo descrito por Hebb (1949), em que a ativação simultânea de neurônios leva a um aumento na força sináptica<sup>5</sup>. Esse exemplo ilustra um dos grandes méritos dos Modelos de Uso: a busca pela integração e intercâmbio com descobertas de outros campos do conhecimento para formulação de modelos linguísticos, que funciona como um fator de coerência e de fortalecimento teóricos.

Evidências mostram que há redundância, não só no uso da linguagem, mas nas unidades de armazenamento e processamento linguístico, contrariando a ideia de que a percepção da fala envolveria a normalização ou eliminação de características e informações não estritamente relevantes para a compreensão da mensagem (cf. JOHNSON; MULLENIX, 1997; PORT, 2007). Estudando a otimização do processamento morfológico por pistas prosódicas utilizadas pelos falantes, Kemps et al. (2005) oferecem evidências de que informações prosódicas relativas a palavras específicas são armazenadas pelo falante e ouvinte. Com base em uma série de experimentos de identificação lexical nas modalidades visual e auditiva, Baayen et al. (2003) concluem que formas de plural de verbos e nomes são armazenadas de forma independente – não somente as irregulares, mas também as regulares são armazenadas. Em testes visando a aplicação à linguagem de um modelo de memória de

---

<sup>3</sup> *Construals* são conceitualizações elaboradas pelo falante ou ouvinte sobre uma situação, envolvendo ajustes focais e construções imagéticas (LANGACKER, 1987, p. 487-488).

<sup>4</sup> Bybee (2010), remetendo a Newell, define *chunks* como unidades da memória formadas por informações reunidas de modo composto, que desempenhariam um papel importante no estabelecimento de uma organização hierárquica na memória e seriam um mecanismo atuante na formação de frases idiomáticas e construções.

traços episódicos, Goldinger (1997) demonstra que informações não linguísticas sobre o falante são retidas pelo ouvinte por um período relativamente amplo. Essas evidências dão suporte à hipótese de que as memórias são ricas, multirrepresentacionais, e de que o léxico é enriquecido, estruturado pela redundância.

A memória linguística é construída por outras memórias – imagéticas, auditivas, motoras –, com as quais interage, e envolve a formação de unidades de conhecimento sobre o mundo em variados níveis de abstração: esquemas. Esquemas desempenham um papel fundamental na percepção: funcionam como sistemas que aceitam informação (*information-accepting systems*): “[...] environmental information is meaningful only insofar as it fits into or relates to schemata that the individual has.” (MOATES; SCHUMACHER, 1980, p. 34). Nos Modelos de Uso, os esquemas – ao contrário das regras, em modelos baseados em regras – não são autônomos ou independentes de instâncias de uso, das quais, pelo contrário, emergem. Pode-se explicar, dessa forma, tanto o processamento mecânico, quanto o analítico e interações entre os dois tipos de mecanismos.

Representações mentais são definidas na forma de esquemas multirrepresentacionais construídos a partir do processamento cognitivo da experiência e englobam informações linguísticas de todos os tipos e também informações não linguísticas (sociais, sensações, etc.); são individuais e também compartilhadas.

Ressalte-se que Modelos de Uso e a Linguística Cognitiva compartilham a ideia de que não há uma separação nítida entre léxico e gramática, essas duas esferas sendo agrupamentos de estruturas simbólicas (LANGACKER, 2000). A linguagem vai muito além de um léxico e um sistema sintático: é melhor compreendida como uma rede de conexões difusas e gradientes, onde não têm lugar distinções definidas e definitivas (BROCCIAS, 2006). Padrões morfológicos e sintáticos são vistos como esquemas (padrões gerais que variam de níveis mais baixos a níveis mais altos em termos de abstração e complexidade) emergentes a partir de ocorrências reais de uso linguístico. Nas abordagens emergentistas (cf. MACWHINNEY, 1999), a estrutura linguística emerge a partir da interação entre processos conhecidos de desenvolvimento.

Finalmente, os Modelos de Uso assumem que “[...] frequency information plays an intrinsic role in the [linguistic] system because it is implicitly encoded by the very nature of the memory system.” (PIERREHUMBERT, 2001a, p. 2). Uma vez que a experiência é

---

5 “The general idea is an old one, that any two cells or systems of cells that are repeatedly active at the same time will tend to become ‘associated’, so that activity in one facilitates activity in the other.” (HEBB, 1949, p. 70).

armazenada diretamente nas memórias linguísticas, a frequência de uso desempenha um papel importante nesses modelos. A frequência de uso como termo geral engloba dois tipos de frequência lexical. A frequência de tipo (*type frequency*) corresponde ao número de itens de um *corpus* que apresentam um determinado padrão (ou tipo), sendo que padrões podem ser palavras, morfemas, sílabas, etc. Já a frequência de ocorrência (*token frequency*) corresponde a quantas vezes um padrão efetivamente ocorre em um *corpus*. Por exemplo, nesta página, até este ponto, o tipo “-mento” apareceu em três itens diferentes (“processamento”, “agrupamentos”, “desenvolvimento”). Ou seja, a frequência de tipo de “-mento” nesta página é três. Além disso, “-mento” ocorreu ao todo quatro vezes nesta página (pois “processamento” aparece duas vezes). Por isso, dizemos que a frequência de ocorrência do tipo “-mento” é quatro. Frequência de tipo e de ocorrência exercem impacto, respectivamente, sobre a produtividade de padrões e sobre a força lexical. Nos Modelos de Uso, a frequência de uso é intrínseca à representação cognitiva das categorias linguísticas.

São bem documentados na literatura os efeitos exercidos pela frequência de uso em diversos aspectos da linguagem humana (cf. BYBEE; HOPPER, 2001). Como exemplo, mencionaremos alguns estudos. Rosenzweig e Postman (1958) demonstraram que a compreensibilidade das palavras aumenta em função de seu tamanho e de sua frequência de uso. Connine et al. (1993) avaliaram a influência da frequência lexical na identificação de fonemas em tarefa envolvendo estímulos ambíguos. Verificaram que os itens ambíguos foram preferencialmente interpretados como a palavra de alta frequência. Pagel et al. (2007), em análise comparativa, identificaram um mecanismo de evolução linguística em que a frequência de uso das palavras seria crucial para os processos de mudança: as línguas começariam a divergir inicialmente em partes menos usadas de seu vocabulário. Além disso, palavras de alta frequência têm tendência maior a exibir propriedades morfológicas “ancestrais”. Algumas palavras mudam o suficientemente devagar de forma a permitir a manutenção de formas análogas por centenas de anos. Nielsen (2011) investigou a influência da frequência lexical e da densidade da vizinhança lexical na identificação de palavras a partir de estímulos auditivos e visuais. Os resultados obtidos mostram que, no paradigma de reconhecimento exclusivamente auditivo, palavras mais frequentes foram reconhecidas melhor, enquanto que no paradigma auditivo-visual, palavras menos frequentes foram reconhecidas melhor. Além disso, palavras pertencentes a vizinhanças lexicais densas foram reconhecidas mais acuradamente que palavras de vizinhanças pouco densas.

Resumindo os conceitos descritos acima, defende-se que as unidades linguísticas de armazenagem e processamento, os exemplares, são palavras e até mesmo grupos de palavras (*chunks*) e colocações. Eventos de uso são armazenados nos exemplares por meio de processos cognitivos gerais de categorização, resultando em grupos de memórias perceptuais ricas, que são conectadas em redes relacionais baseadas em similaridade fonética e semântica. Tais redes, a partir das quais emergem esquemas gerais (morfológicos, sintáticos, etc.), operam com base em parâmetros probabilísticos, apresentando distribuição estocástica (PIERREHUMBERT, 2001b). A FIG. 3 abaixo mostra uma rede de exemplares e a emergência de um padrão morfológico de passado no inglês.

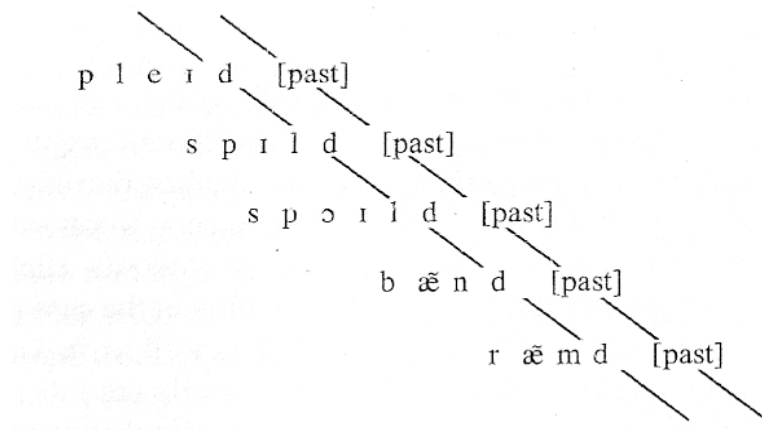


FIGURA 3 - Rede de exemplares e emergência de padrão morfológico

FONTE: BYBEE, 2010, p. 23.

Como se pode ver, na FIG. 3, os exemplares encontram-se conectados por relações fonológicas e semânticas, cuja recorrência ao longo de inúmeros exemplares leva à emergência de uma estrutura abstrata, que corresponde ao esquema regular de formação de passado pelo acréscimo de [d] à forma presente de um verbo.

Os princípios básicos dos Modelos de Uso e da Linguística Cognitiva descritos acima têm sido empregados com sucesso na análise de fenômenos do plano fonético/fonológico que representam problemas para as análises linguísticas tradicionais, tais como o conhecimento que os falantes possuem acerca de detalhes fonéticos finos, efeitos de prototipicidade e efeitos de frequência (BYBEE; HOPPER, 2001). Pretendemos, neste estudo, contribuir com o desenvolvimento de Modelos de Uso e da Linguística Cognitiva, aplicando seus conceitos básicos na descrição de um tema ainda pouco explorado nesta perspectiva, os mecanismos de proeminência prosódica. Na análise do acento lexical a ser desenvolvida no

cap. 5, defenderemos que tanto mecanismos gerais abstratos quanto especificidades de unidades particulares de uso linguístico entram em jogo na dinâmica acentual do PB. Essa proposta se compromete com a ideia de que a linguagem emerge a partir de mecanismos cognitivos gerais de categorização e processamento, em que especificidade e generalidade não são excludentes, mas antes coexistem e se complementam.

### 3 Acento e produção de fala

Este capítulo apresenta abordagens teóricas e experimentais do acento sob um ponto de vista da produção da fala. Primeiramente, iremos delimitar o fenômeno do acento lexical, apresentando observações de ordem teórica e terminológica. Traçaremos então um panorama histórico sobre como evoluiu a concepção do acento em algumas teorias linguísticas, do estruturalismo aos modelos gestuais, passando pelo gerativismo e modelos fonológicos não lineares. Em seguida, serão descritas as propriedades acústicas e articulatórias geralmente associadas ao fenômeno de proeminência acentual. Na subseção seguinte, resumiremos os aspectos consensuais e os principais debates enfrentados pelas teorias fonológicas sobre o acento lexical no português brasileiro. Por fim, apresentaremos alguns estudos acústicos do acento no português brasileiro.

#### 3.1 Delimitação do fenômeno acentual

Tradicionalmente, os estudos sobre a dimensão sonora da fala humana são divididos em Fonética e Fonologia. Enquanto a primeira se dedica ao detalhamento dos aspectos físicos dos sons da fala – em termos de suas propriedades acústicas, aerodinâmicas, articulatórias, auditivas –, a segunda se volta para o funcionamento e a organização de tais sons – em termos de estruturas e mecanismos cognitivos subjacentes. Trata-se, portanto, de uma oposição quanto à perspectiva de análise adotada<sup>6</sup>. A área da Prosódia, que comporta tanto uma abordagem do tipo fonético quanto fonológico, se define antes pela especificidade de seu tema: dedica-se ao estudo de fenômenos não segmentais, geralmente atribuídos à “voz”, como ritmo, entonação e acento. Ritmo (*rhythm*) se refere aos padrões que surgem na fala “[...] out of the periodic recurrence of some sort of movement, producing an expectation that the regularity of succession will continue.” (ABERCROMBIE, 1968, p. 96). A entonação

---

<sup>6</sup> A divisão da dimensão sonora em Fonética e Fonologia é controversa tanto em sua separação meramente metodológica, em vista da interferência entre os dois níveis de análise (não raro o fonólogo toma decisões com base em detalhes físicos dos sons e o foneticista, com base em seu funcionamento), quanto em sua separação orgânica respectivamente no módulo de implementação e no módulo de planejamento da linguagem, sendo questionada a oposição entre a capacidade da linguagem e seu uso (cf. cap. 2).

(*intonation*) é um dos aspectos que contribuem com a organização do ritmo da língua e diz respeito ao uso da variação da altura para transmitir “[...] sentence-level pragmatic meanings in a linguistically structured way.” (LADD, 2008, p. 4). A definição de acento (*stress*), fenômeno que também contribui para as características rítmicas de uma língua e tema central deste trabalho, se abre a variadas perspectivas. De um lado, há definições que enfocam a interação do acento com o plano entonacional: “Stress is an abstract positional property of linguistic units which denotes the position at which a non-contrastive accent [i.e. intonational prominence] can occur. The stress position is determined by the structure of words and phrases.” (HIRST; DI CRISTO, 1998, p. 86). Por outro lado, em uma perspectiva fonética, como a de Hogg e McCully (1987, p. 1-2), “[...] stress is associated with an increase in respiratory activity on the part of the speaker which causes an increase in any one or other of the following: vowel length, pitch, loudness.” Note-se que a definição de Hog e McCully relaciona um mecanismo aerodinâmico com grandezas perceptuais (*length, pitch e loudness*, que preferimos não traduzir, para fugir a problemas terminológicos), definidas como correlatos das grandezas acústicas duração, frequência fundamental (F0) e intensidade. No capítulo 4, que trata da percepção de padrões acústicos, retomaremos a relação entre os três pares de propriedades mencionados acima.

Para alguns autores, ainda, acento pode ser definido simplesmente como proeminência silábica:

In the past, the word ‘stress’ has been used in different and confusing ways. It has sometimes been used simply to refer to syllables (or vowels) made prominent for linguistic purposes, either in words or in sentences. But stress has also often been used to mean ‘breath-force or loudness’ the implication being that this is the principal means whereby syllables are made prominent. This second type of usage is misguided since [...] loudness generally plays a minor role in producing prominences. [...] I shall use the term STRESS to mean ‘prominence’, however such prominence is achieved. (CRUTTENDEN, 1997, p. 13).

A definição de “acento” ou “acento lexical” que será usada neste estudo buscará integrar as três perspectivas apresentadas acima. Contudo, Hyman (1977) aponta para a necessidade de se conhecer mais a respeito do acento. Anos se passaram e resta ainda grande controvérsia a respeito da própria definição conceitual e terminológica do fenômeno, sem que se tenha chegado ainda a uma posição consensual sobre sua natureza, como se vê em texto relativamente recente do mesmo autor (HYMAN, 2006). Por exemplo, ainda que a oposição

entre acento e tom seja de um modo geral aceita e definida com alguma facilidade, surgem problemas ao se diferenciar acento no que Hyman chama de *stress-accent* e *pitch-accent*.

Nota-se que a Prosódia enfrenta ainda muitos problemas quanto a definições terminológicas, como retomaremos adiante, em especial problemas referentes à tradução dos conceitos prosódicos e à multiplicidade de nomenclaturas. Para fins de clareza, iremos apresentar o termo original em inglês sempre que um novo conceito for introduzido no texto. Destacamos ainda que será evitado o termo “tonicidade” como referência ao fenômeno do acento, uma vez que etimologicamente “tônico” pode passar uma falsa ideia de que se trata de uma marcação acentual baseada em F0. Como veremos mais adiante, o acento apresenta simultaneamente mais de um correlato acústico possível. Exceção será feita para denominar as sílabas átonas em dois tipos: as que antecedem a sílaba acentuada (“pretônicas”) ou que a sucedem (“postônicas”), pela falta de termos substitutivos eficientes. Além disso, o termo “acento” que será utilizado ao longo deste estudo corresponde à tradução de *stress*, e não de *accent*, sendo o último geralmente empregado com referência a características da pronúncia de um indivíduo que o identificam com sua comunidade de fala ou ainda à ênfase conferida a palavra ou sílaba em níveis prosódicos superiores, como o sintagma, a frase ou o enunciado (CRYSTAL, 1995).

Em algumas línguas, o acento lexical pode recair sobre mais de uma sílaba na mesma palavra, havendo uma gradação no nível de proeminência em cada uma das sílabas acentuadas. Nessa caso, diz-se que esta língua apresenta acento primário (que incide sobre a sílaba mais proeminente) e também secundário (incidente sobre as sílabas com menor grau de proeminência).

[...] in any word [...] one and only one vowel receives the strongest stress of the word. We should call this stress *primary stress*. [...] the vowels in any constituent which are stressed but do not receive the strongest stress of the constituent are said to bear *secondary stress*. [...] Only vowels which are initially assigned primary stress can, in some larger context, bear secondary stress (HOGG; MCCULLY, 1987, p. 6).

Iremos restringir nossa análise somente ao acento lexical primário do PB, visto que o acento secundário é um fenômeno complexo, que mereceria um estudo à parte. Abordagens aprofundadas do assunto podem ser encontradas e.g. em Collischonn (1994) e Arantes (2010a). No presente trabalho, tanto “acento lexical” quanto “acento” serão empregados como equivalentes a “acento lexical primário”.

Na perspectiva que Abercrombie (1968) denomina de Fonética Clássica, os componentes acústicos da fala foram divididos em segmentais e não segmentais (ou prosódicos), sendo os últimos subdivididos em qualidade vocal e dinâmica vocal.

The strand consisting of the segmental features of an utterance is made up of complex auditory qualities which are in fairly rapid fluctuation, reflecting the rapid succession of movements of the articulators. The strand of voice dynamics also consists of features which fluctuate in auditory quality, but considerably more slowly. Many of them are linked to the syllable- and stress-producing processes [...], and they are closely related to [...] aspects of sound [...] such as pitch loudness, tempo ad rhythm. The strand consisting of features of voice quality in contrast to the two preceding, has a quasi-permanent character: it remains constant over relatively long stretches of time, and fluctuation here is much less apparent. (ABERCROMBIE, 1968, p. 90).

Os fenômenos prosódicos são, via de regra, realizados por meio de três parâmetros acústicos básicos – F0, intensidade e duração –, podendo haver sobreposição entre os três parâmetros na implementação de diferentes fenômenos (LADEFOGED, 1996; LAVER, 1994). Em uma mesma língua, por exemplo, acento e entonação, ainda que implementados independentemente, podem se valer de mecanismos acústico-articulatórios similares. Além disso, a interação entre os mecanismos acústicos de implementação do acento apresenta especificidades em cada língua. Por essa razão, uma definição puramente fonética do acento, como “[t]he use of extra respiratory energy during a syllable” (LADEFOGED, 1982, p. 285), não caracteriza todo o fenômeno. Mostra-se necessário descrevê-lo considerando também suas características funcionais. O presente estudo busca oferecer contribuições para a descrição do acento sob um ponto de vista funcional e experimental.

## 3.2 O estudo do acento ao longo do desenvolvimento histórico das teorias fonológicas

### 3.2.1 Estruturalismo

A tradição estruturalista inaugurou a descrição dos sons das línguas do mundo por meio da combinação de propriedades básicas, os traços. Tais propriedades foram

primeiramente elaboradas nos estudos fonéticos, como forma de descrever os segmentos sonoros físicos (fones) de acordo com suas características articatórias. Foram então reelaboradas em uma perspectiva fonológica e passaram a contemplar também características acústicas, com fins de definir os segmentos sonoros abstratos (fonemas). Construiu-se um sistema, que se supunha universal, de traços distintivos, definidos como as unidades mínimas no plano sonoro<sup>7</sup>. A partir dos traços, seriam obtidas, combinatorialmente, todas as estruturas sonoras da língua: fonemas, sílabas, palavras.

The distinctive features are aligned into simultaneous bundles called phonemes; phonemes are concatenated into sequences; the elementary pattern underlying any grouping of phonemes is the SYLLABLE. The phonemic structure of the syllable is determined by a set of rules, and any sequence is based on the regular recurrence of this constructive model. A FREE FORM (a sequence, separable by means of pauses) must contain an integral number of syllables. [...] The pivotal principle of syllable structure is the contrast of successive features within the syllable. One part of the syllable stands out against the others. It is mainly the contrast vowel vs. consonant which is used to render one part of the syllable more prominent. (JAKOBSON; HALLE, 1971, p. 31).

A seguir o QUADRO 1 esquematiza, a título de ilustração, a tipologia de traços sonoros proposta por Jakobson e Halle (1971)<sup>8</sup>. Note-se que há traços que se referem a propriedades acústicas, articatórias e funcionais dos segmentos.

---

7 Para diferentes propostas de sistemas de traços fonológicos, cf. Kenstowicz (1993).

8 Os nomes dos traços não foram traduzidos, para evitar problemas terminológicos não pertinentes a este estudo. Os pares de traços agrupados em cada categoria – força, quantidade e tom – correspondem, de cima para baixo, ao nível intersilábico e intrasilábico.

		<b>Traços</b>				
Distintivos	Prosódicos	força	<i>stress</i> <i>stosston</i>			
		quantidade	<i>length</i> <i>contact</i>			
		tom	<i>level</i> <i>modulation</i>			
	Inerentes	sonoridade		<i>vocalic/non-vocalic</i> <i>consonantal/non-consonantal</i> <i>nasal/oral</i> <i>compact/diffuse</i> <i>abrupt/continuant</i> <i>strident/non-strident</i> <i>checked/unchecked</i> <i>voice/voiceless</i>		
			“protensidade”	<i>tense/lax</i>		
			tonalidade	<i>grave/acute</i> <i>flat/non-flat</i> <i>sharp/non-sharp</i>		
			Configurativos	Culminativo		
				Demarcativo		
Redundantes						

QUADRO 1 - Sistema de traços fonológicos proposto em Jakobson e Halle (1971)

Traços distintivos do tipo prosódico seriam propriedades exclusivas de segmentos que exercem função de núcleo da sílaba (geralmente vogais), enquanto que os traços do tipo inerente abrangeriam todos os tipos de segmentos, independente de seu papel na estrutura silábica. Os autores agrupam os traços prosódicos nas categorias de força, quantidade e tom, correspondentes, no plano perceptual, respectivamente a *loudness*, duração relativa e *pitch*, e, no plano acústico, a intensidade (amplitude), duração absoluta (tempo) e frequência.

Os autores argumentam que, ao contrário dos traços inerentes, tidos como absolutos, os traços prosódicos implicam no estabelecimento de contraste entre variáveis em determinado domínio: por exemplo, uma sílaba é considerada mais proeminente em comparação a outra sílaba da mesma palavra.

A prosodic feature involves two coordinates: on the one hand, polar terms such as high or low register, rising and falling pitch, or long and short, all may appear, *ceteris paribus*, in the same position within the sequence; thus

the speaker selectively uses and the listener selectively apprehends one of the two alternatives and identifies the chosen alternative in relation to the rejected one. [...] On the other hand, both polar terms are fully recognizable only when both of them are present in the given sequence, so that the speaker effects and the listener perceives their contrast. Thus both alternative of a prosodic feature co-exist in the code as two terms of an opposition and, moreover, co-occur and produce a contrast within the message. (JAKOBSON; HALLE, 1971, p. 37).

Traços prosódicos são, portanto, propriedades sempre relativas e, conseqüentemente, variáveis em termos de sua magnitude absoluta, seja entre falantes diferentes ou entre locuções de um mesmo falante.

Outro aspecto destacado por Jakobson e Halle diz respeito à função do acento, especificada por meio dos traços configurativos que poderiam interagir com o traço distintivo *stress*. A função culminativa diz respeito à presença de apenas um pico máximo de proeminência local (seja em uma palavra, geralmente de conteúdo, ou em um grupo formado por palavras funcionais anexadas a palavras de conteúdo<sup>9</sup>), que serve como âncora para a estrutura prosódica (PIERREHUMBERT, 1980; PIERREHUMBERT; BECKMAN, 1988). Já a função demarcativa do acento se refere à sinalização de fronteiras gramaticais, seja de morfemas, palavras, ou grupos de palavras.<sup>10</sup>

Wherever there is a contrast of stressed and unstressed syllables, stress is always used as a configurative, namely culminative feature, whereas length never assumes this function. The culminative function of stress is regularly combined either with the other variety of configurative functions, demarcation, or with the distinctive function. Languages in which both length and stress appear as mutually independent distinctive features are quite exceptional, and if the stress is distinctive, it is frequently supplemented by redundant length. The observation of force and quantity features in their intersyllabic variety seem to indicate that the prosodic distinctive features utilizing intensity and those utilizing time tend to merge. (JAKOBSON; HALLE, 1971, p. 36).

Como se pode ver, os autores ressaltam que, em várias línguas, como consequência da coocorrência redundante dos parâmetros acústicos de intensidade e duração na realização do acento, pode haver fusão entre essas propriedades. Como retomaremos na seção seguinte, o acento lexical como o do português mostra-se um caso complexo para a caracterização das

9 Compare, e.g., a presença de duas proeminências acentuais nos sintagmas “**bela casa**” e apenas uma em “a **ca**sa”. (Ao longo deste estudo, quando relevante, sílabas acentuadas serão marcadas por negrito e representação fonética dos sons será dada entre colchetes: “[ ]”).

10 No português, o acento lexical acontece sempre em uma das três últimas sílabas da palavra (e.g. **cará**, **cara**, **máscara**) ou na quarta última, em casos específicos de epêntese e próclise (e.g. **rit**[i]mico e **chamáv**amo-nos).

unidades sonoras por meio de traços: é ao mesmo tempo um traço do tipo distintivo prosódico (*stress e length*), do tipo configurativo (culminativo e demarcativo) e do tipo redundante (dada a fusão entre os traços prosódicos). Além disso, sua função configurativa como âncora prosódica e delimitador de palavras é um indicativo de que sua abrangência extrapola o domínio de segmentos individuais, como veremos adiante.

### 3.2.2 Gerativismo e modelos não lineares

O gerativismo não apresentou mudanças significativas com relação ao estruturalismo quanto à concepção de acento. Chomsky e Halle (1968), considerando a fonologia do inglês, definem o acento como um traço [+– stress] atribuído à vogal por meio de uma regra específica da língua (*Nuclear Stress Rule*) e que depende da estrutura sintática. No entanto, a abordagem baseada em traços, com atribuição do acento por meio de algoritmos específicos, como a defendida pela fonologia gerativista, logo se deparou com problemas em lidar com características prosódicas da fala.

Alguns dos modelos fonológicos que surgiram nas décadas de 70 e 80, ditos Modelos Não Lineares<sup>11</sup>, buscaram solucionar problemas de interação dos segmentos com a prosódia. Uma das grandes inovações dessa época foi a proposta de interpretar o plano sonoro em uma estrutura multi-linear. A Fonologia Autossegmental desenvolveu a ideia de que os fenômenos prosódicos seriam suprasegmentos (GOLDSMITH, 1976) – i.e. ocorreriam em um plano hierárquico superior ao plano segmental, uma vez que podem operar sobre uma cadeia de segmentos. O modelo autossegmental ganhou elaboração teórica com Geometria de Traços (CLEMENTS, 1985), em que a matriz de traços definidores de um segmento passa a ser vista não simplesmente como uma lista desordenada de atributos independentes. Pelo contrário, a Geometria de Traços sugere que os traços podem ser agrupados hierarquicamente em nós referentes a pontos articulatórios. Os traços seriam atados a um nó de raiz, sendo este dominado por uma camada temporal, representando a duração do segmento. Segmentos são associados aos constituintes silábicos – *onset*, núcleo, coda –, que, por sua vez, se vinculam à

---

<sup>11</sup> Os modelos fonológicos chamados de não lineares se comprometem com os axiomas fundamentais da Gramática Gerativa, buscando contribuir com a adequabilidade desse programa de pesquisa na área do componente sonoro.

sílaba e ao plano de atribuição do acento. Uma das contribuições mais significativas da Fonologia Autossegmental para o estudo do acento, ao lado da noção de não linearidade – i.e., multidimensionalidade – dos suprasegmentos, está relacionada à noção da sílaba como unidade de *status* teórico.

Outro modelo não linear que paralelamente ofereceu avanços para o estudo do acento foi a Fonologia Métrica, que, voltando-se para a estrutura rítmica de palavras e sentenças, conferiu um tratamento unificado ao acento lexical e frasal, como fenômenos da mesma natureza, diferenciando-se apenas por se aplicar a domínios diferentes: “[...] stress patterns reflect, to some extent at least, relations of prominence between constituents. [...] Thus stress assignment in metrical phonology is a matter of denoting the relative prominence of sister constituent.” (HOGG; MCCULLY, 1987, p. 64-65). Liberman e Prince (1977) defendem que o acento não é uma propriedade de segmentos individuais, ou sílabas, como proposto pelo estruturalismo e gerativismo, mas antes reflete uma estruturação rítmica hierárquica responsável por organizar as sílabas, palavras e sentenças de um enunciado. Os autores apontam inconsistências na abordagem do acento como um traço distintivo, argumentando que o traço [*stress*] não se comporta como os outros traços do sistema: os valores do traço [*stress*] são n-ários; a possibilidade de níveis diferentes de acento dentro de um mesmo domínio leva à necessidade de regras de acento secundário, terciário, etc., no nível sintagmático, enquanto os outros traços fonéticos são definidos paradigmaticamente; regras acentuais não são locais, como as regras segmentais. Tais problemas seriam solucionados pela abordagem métrica, com o estabelecimento de uma organização binária do acento na estrutura arbórea de constituintes métricos, que definiria relações sintagmáticas, de nível local.

Com a repercussão da abordagem não linear, ênfase foi dada à descrição dos fenômenos prosódicos em termos de seu domínio de ocorrência: há padrões que ocorrem no nível da palavra (*word-level prosody*, como o acento e o tom) e padrões que ocorrem no nível da frase (*phrase-level prosody*, como a entonação), e a organização de tais domínios obedece a princípios hierárquicos estritos, como focado na Fonologia Prosódica de Selkirk (1980) e Nespore e Vogel (1986). Ênfase foi dada à descrição e demonstração da relevância de constituintes prosódicos (HALLE; VERGNAUD, 1987; MCCARTHY; PRINCE, 1999).

Muito se debateu nas teorias formalistas não lineares sobre aspectos paramétricos do acento, específicos de cada língua, especialmente quanto a: número de constituintes por pé; pé com núcleo à esquerda ou à direita; iteratividade e direcionalidade na divisão dos pés;

sensibilidade ao peso silábico (cf. HAYES, 1995), com objetivo de desenvolver uma estrutura de princípios prosódicos universais. Alguns desses debates ganharam nova roupagem com o desenvolvimento da Teoria da Otimalidade (PRINCE; SMOLENSKY, 2002), que apresenta uma nova perspectiva de gramática, rompendo com a noção de regras e representação subjacente. As regras dão lugar ao ranqueamento de restrições violáveis. Parte-se do *output* (que corresponderia à forma de superfície na gerativa) e, através da interação de restrições, escolhe-se o candidato ótimo, que será o *input* (o que era a forma subjacente na gerativa).

### 3.2.3 Modelos gestuais

A descrição e modelagem da articulação dos sons da fala ganhou importância crescente nos estudos linguísticos, ao longo das três últimas décadas. A fala passou a ser vista como produto da multimodalidade, dando origem a diversos modelos gestuais em fonologia (e.g. Fonologia Articulatória (BROWMAN; GOLDSTEIN, 1986), *Task Dynamics* (SALTZMAN; MUNHALL, 1989)). Para tais modelos, a fala é construída a partir dos gestos, padrões articulatórios gerados pela movimentação coordenada dos articuladores da fala. Gestos são unidades representacionais definidas dinamicamente, nas dimensões espacial e temporal (GAFOS, 2002). Os modelos gestuais foram primeiramente desenvolvidos para lidar com fenômenos segmentais, mas, na última década, passaram a dispor também de abordagens da estrutura prosódica (BYRD; CHOI, 2010; SALTZMAN et al. 2008). Em comparação com as outras teorias fonológicas, o PB conta com poucos estudos baseados em modelos gestuais. Dos que tratam do acento lexical, podemos mencionar Albano (2001), que distingue a potencialidade de proeminência lexical do acento propriamente dito. A autora atribui a potencialidade de proeminência à magnitude do gesto vocálico, assumindo que a representação das palavras conteria informações sobre as vogais que são acentuáveis: “a vogal final ou pré-final da palavra é lexicalmente reduzida exceto se a fronteira vocabular coincide com certas fronteiras morfológicas” (ALBANO, 2001, p. 150-151). A acentuação se realizaria, então, pela atribuição de proeminência fonética às vogais acentuáveis, o que coloca a abordagem de Albano em posição intermediária entre a hipótese de armazenagem lexical e de implementação por algoritmos.

### 3.3 Aspectos físicos da proeminência acentual

Esta subseção explora a dimensão física do acento, em termos de suas características acústicas e articulatórias. Para tal, é conveniente antes apresentar algumas definições básicas da análise acústica da fala, baseadas em Handel (1993), Stevens (2000), Kent e Read (1992), aos quais remetemos o leitor para uma visão aprofundada sobre o assunto.

O som surge de vibrações formadas por diferenças de pressão que se propagam sob a forma de uma onda. Na Física, as medidas básicas empregadas para descrever uma onda, relevantes para este estudo, são duração, período, amplitude, como representado na FIG. 4. A duração é uma medida intuitiva e corresponde ao tempo transcorrido do início ao fim de um evento. Em eventos envolvendo um padrão que se repete ao longo do tempo, o período é o tempo transcorrido até que o padrão volte seu valor inicial. A amplitude<sup>12</sup> corresponde ao valor máximo que o padrão atinge. A partir do período e da amplitude, podemos obter as duas outras medidas que, ao lado da duração, são tradicionalmente empregadas na análise da fala: frequência<sup>13</sup> e intensidade. A frequência é o inverso do período e corresponde ao número de repetições de um padrão em determinado espaço de tempo. A relação entre amplitude e intensidade não é tão simples e a intensidade não é tão facilmente visualizável como as demais medidas, mas, para os presentes propósitos, basta dizer que a intensidade é proporcional ao quadrado da amplitude.

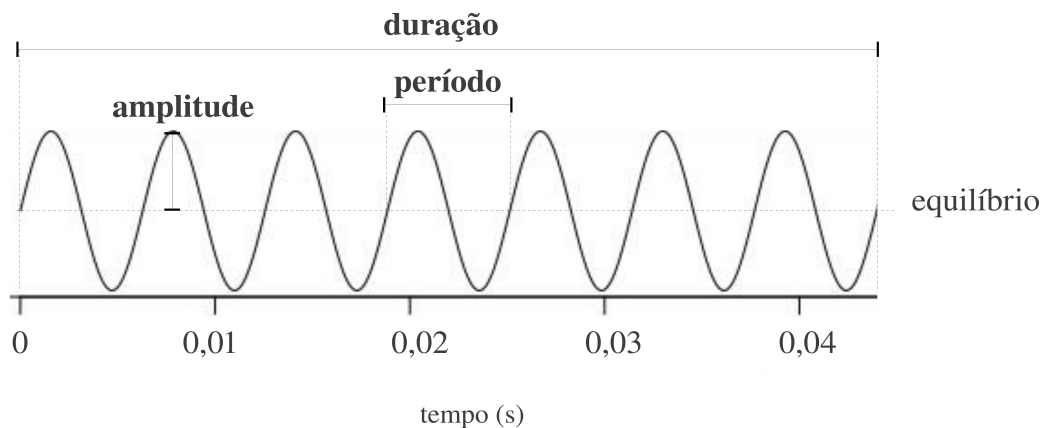


FIGURA 4 - Medidas básicas realizadas em uma onda

<sup>12</sup> Amplitude de pico.

<sup>13</sup> Frequência temporal, que será referida apenas como frequência. Em outra parte deste trabalho, iremos introduzir uma medida de contagem lexical, também chamada de frequência, mas que, para evitar confusões, será especificada como frequência lexical (de ocorrência ou de tipo).

Note-se que a FIG. 4 corresponde à representação de um som formado por uma única frequência. Os sons da fala, por outro lado, são complexos, pois podem ser decompostos em várias frequências de vibração. A frequência de valor mais baixo é chamada de frequência fundamental e as frequências superiores são chamadas de harmônicos e seus valores são múltiplos inteiros da fundamental. Nos sons da fala, a frequência de vibração das cordas vocais corresponde à frequência fundamental, também chamada de F0.

A produção de um som requer um posicionamento característico dos articuladores no trato, resultando em uma cavidade com modos de vibração particulares. Tais modos de vibração realizam uma amplificação seletiva de certas regiões de frequências do sinal inicial. As frequências de ressonância do trato vocal, também chamadas de formantes, correspondem a regiões espectrais com alta concentração de energia, formadas pela amplificação seletiva do sinal glótico. Em sons vocálicos, os formantes são infinitos em número, mas apenas os três primeiros (F1, F2, F3) são considerados na análise fonética, pois são suficientes para a identificação de cada som.

Um modelo clássico em acústica para descrição da fala, conhecido como modelo de fonte-filtro, se baseia no princípio de que os sons da fala consistem em uma onda sonora que é originada de uma fonte de vibração e modificada ao passar por uma cavidade de ressonância. Tal modelo foi amplamente explorado por Fant na Teoria Acústica da Produção da Fala (FANT, 1970), que, ao lado de outros grandes foneticistas de sua época, foi responsável por consolidar métodos experimentais de análise acústica da fala e por descrever os mecanismos acústicos e articulatórios envolvidos na produção da fala. O modelo fonte-filtro descreve de forma simplificada os processos envolvidos na produção dos sons da fala, procurando estabelecer uma interpretação causal entre eventos acústicos e articulatórios, de forma a ser possível prever certos resultados sonoros a partir de configurações gestuais e vice-versa.

Na descrição e modelagem física dos fenômenos prosódicos, são geralmente considerados, além dos aspectos acústicos, aspectos aerodinâmicos e da articulação no nível glótico. Procura-se determinar os mecanismos geradores das propriedades acústicas do acento e da entonação, em especial, os correlatos articulatórios da intensidade e da frequência.

Fant (1981) aponta duas estratégias básicas envolvidas no aumento da intensidade: (1) aumento na pressão subglótica; (2) adução das cordas vocais devido à compressão medial. F0 e intensidade se correlacionam positivamente com a pressão subglótica, cujo aumento pode resultar da ação do sistema respiratório ou do aumento na resistência da glote à corrente

de ar (FANT, 1981; LADEFOGED; MCKINNEY, 1963; ROTHENBERG, 1973). Contudo, independentemente da ação da pressão subglótica, F0 também é influenciada pela tensão das cordas vocais. Dessa forma, a articulação na região laringea poderia desempenhar algum papel na diferenciação de sílabas acentuadas e não acentuadas. Baseando-nos nessa possível relação acústico-articulatória, optamos por utilizar a eletroglotografia em nosso experimento, que será apresentado no cap. 6. Aumentos em F0 e, indiretamente, intensidade poderiam ser identificados na forma de alterações do sinal glótico.

No inglês, Jong (1995) avaliou os correlatos do acento na movimentação oral e da cabeça por meio de imagens de raio-X e obteve os seguintes resultados: há diferenças individuais nas estratégias de articulação do acento; mandíbula, lábio e língua apresentam interação dependente do sujeito; o acento exerce efeito sobre a posterioridade de vogais e o ponto de articulação de consoantes. Hadar (1991) encontrou correlação da movimentação da cabeça com padrões rítmicos da fala. Evidências de relações entre a acentuação e movimentação da face motivaram a adoção, em nosso experimento de produção, de um equipamento de captura de movimentos, o Optotrak, como detalharemos no cap. 6.

Nesta seção, procuramos apresentar resultados de estudos sobre a fala que mostram a importância de informações acústicas e articulatórias para a compreensão de fenômenos linguísticos e nos quais nos basearemos para analisar o acento. Apontamos evidências de que propriedades acústicas envolvidas na produção do acento poderiam ser relacionadas à articulação no nível glótico e no nível da face e cabeça, o que nos motivou a utilizar em nossa coleta de dados um eletroglotógrafo e o Optotrak. Por restrições de tempo, os dados coletados com o Optotrak serão analisados em um trabalho futuro.

Na seção seguinte passaremos a considerar estudos que tratam especificamente do acento no português, apontando suas contribuições para a descrição do fenômeno, bem como aspectos específicos decorrentes do modelo teórico adotado em cada estudo.

### 3.4 Estudos sobre o acento lexical no português

O acento lexical, no PB, se restringe a uma janela de três sílabas, ao fim da palavra, formando oxítonas, paroxítonas ou proparoxítonas (cf. BISOL, 1999; CAMARA JÚNIOR,

1970). Essa janela seria expandida para a quarta sílaba apenas em caso de epêntese (e.g. ['tɛkɪniku], *técnico*, vogal epentética sublinhada), no sistema não-verbal, e de procliticização (e.g. *falávamos-te*), no sistema verbal.

Palavras paroxítonas prevalecem numericamente sobre proparoxítonas e oxítonas, ou seja, dentre os três padrões acentuais possíveis na língua, o paroxítono é o mais frequente (ALBANO, 2001). Este padrão seria responsável por imprimir o ritmo característico do português e do espanhol, em comparação com o acento predominantemente proparoxítono do italiano e exclusivamente oxítono do francês (CAMARA JÚNIOR, 1970, p. 55).

Como sugere Câmara Júnior (1970, p. 54), no PB, o acento tem função demarcativa, servindo para delimitar os vocábulos fonológicos, culminativa, por servir como âncora para a entonação, e distintiva, sendo responsável pelo contraste lexical em palavras como *sábia*, *sabia* e *sabiá*. Além disso, como se pode depreender dos exemplos anteriores, o acento tem também a função de contrastar classes lexicais (LEE, 1995, p. 141), o que recebe suporte do fato de não ser fácil encontrar pares mínimos de acento dentro da mesma classe de palavras – como, por exemplo, *cara* x *cará*, no caso dos nomes, e *comeram* x *comerão*, no caso dos verbos –, fato que geralmente não é explorado pelas análises do acento no PB.

São vários os estudos que tratam da acentuação no PB (cf. Araújo (2007) para um apanhado de várias dessas análises). Apesar de estarem fundamentados em diferentes modelos teóricos, a descrição dos fatos linguísticos neles apresentada é razoavelmente equivalente. A maioria deles aponta para quatro aspectos fundamentais do acento:

- a) tem extensão abrangente, alcançando o domínio da palavra e grupos de palavras;
- b) envolve o estabelecimento de proeminência relativa contrastiva dentro do domínio de ocorrência;
- c) tem função de delimitar palavras, pois ocorre sempre em uma janela de três sílabas e interage com processos fonológicos como o de lenição silábica;
- d) interage com proeminências em outros níveis prosódicos, servindo como apoio para o ritmo e para a entonação.

Nos estudos sobre o acento lexical no português brasileiro, um dos pontos mais discutidos é em que medida ele é (1) especificado na estrutura lexical (“marcado no léxico”), (2) implementado por algoritmo de atribuição de acento ou (2) sensível ao peso silábico. Serão apresentados a seguir estudos que exemplificam cada posicionamento.

Camara Júnior (1970, p. 55), comprometido com a tradição estruturalista, defende que o acento no português é lexicalmente especificado, sendo que “sua posição não depende da estrutura fonêmica do vocábulo”.

Leite (1974) propõe uma análise gerativista do acento no PB. Avalia propostas anteriores, entre elas a de Camara Júnior (1970), e conclui que “stress and other phonetic properties are not independent phonetic characteristics of an utterance, dissociated from grammatical and syntactical considerations. They are the result of the application of a set of rules.” (LEITE, 1974, p. 26). Propõe que a informação sobre o acento não estaria presente no léxico, mas seria atribuída à cadeia sonora por meio de regra específica. Ainda, uma regra levaria em conta a ocorrência de determinados sufixos, todas elas sujeitas a condições específicas de aplicação. Em sua análise, oxítonas terminadas em vogal são marcadas no léxico para o acento. A autora conclui ser necessário postular regras diferentes para o sistema verbal e nominal, uma vez que a regra de atribuição de acento se aplicaria diferentemente em cada um dos sistemas.

Baseando-se na Fonologia Métrica, Bisol (1992, p. 69) apresenta a seguinte regra para atribuição do acento no PB que, ao contrário da regra gerativista de Leite (1974), faz referência a constituintes silábicos<sup>14</sup>:

Regra do Acento Primário

Domínio: a palavra

- i. Atribua um asterisco (\*) à sílaba pesada final, i.e, sílaba de rima ramificada.
- ii. Nos demais casos, forme um constituinte binário (não iterativamente) com proeminência à esquerda, de tipo (\* .), junto à borda direita da palavra.

Esta regra básica de atribuição de acento irá acentuar sílabas pesadas<sup>15</sup> finais (como em *amor*) e, na ausência destas, acentuar a primeira sílaba do pé<sup>16</sup> localizado junto à borda direita da palavra (como em *travessura*, que, dividida em pés, (trave)(ssura), tem o primeiro pé mais à direita, (*ssura*), com a estrutura (\* .)). Portanto, a análise de Bisol (1992) considera a estrutura silábica determinante para a localização do acento e defende que o pé básico no

---

14 A sílaba apresenta uma estrutura hierárquica interna composta por *onset* (ou ataque), seguido de núcleo e coda, sendo estes três elementos chamados de constituintes silábicos. O *onset* é ocupado pelas consoantes iniciais e não exerceria influência na atribuição do acento. O núcleo é o centro da sílaba, componente obrigatório, equivalendo ao pico de sonoridade que a caracteriza, sendo ocupado por vogais, semi-vogais e, em algumas línguas, certas consoantes contínuas. A coda é ocupada por consoantes pós-nucleares. Núcleo e coda podem ser agrupados em um constituinte superior, a rima. (cf. HOGG; MCCULLY, 1987, p. 36-38).

português seria troqueu – representado como (\* ). A autora tenta dar conta das exceções à regra básica recorrendo à extrametricidade<sup>17</sup> (para palavras com sílaba pesada final átona (e.g. *lúpis*) e para todas as proparoxítonas) e à presença de uma consoante final abstrata, ou “invisível” (para oxítonas terminadas em sílaba leve (como em *jacaré*)). Além disso, explica a diferença entre verbos e nomes em termos da natureza da regra de atribuição de acento: cíclica, nos não-verbos, não-cíclica, nos verbos.

Das abordagens do acento no PB que se valem da Teoria da Otimalidade, podemos mencionar Wetzels (2006). O autor sugere que o sistema acentual no PB é misto: nos verbos, o acento tem base morfológica e é sensível às categorias de tempo, enquanto que, nos não-verbos, o acento tem base prosódica e é sensível ao peso silábico. O autor assume que o pé, no PB, tanto em verbos quanto em não-verbos, e mesmo nos casos excepcionais, tem cabeça à esquerda e é alinhado com a borda direita da palavra. Analisando as formas verbais, propõe três restrições para identificar a localização da cabeça do pé, de acordo com o tempo verbal (FOOT HEAD PRESENT, FOOT HEAD PAST, FOOT HEAD FUTURE). Em seguida, passando aos não-verbos, apresenta seis restrições que definiriam as propriedades básicas do acento no PB, ranqueando-as da seguinte forma (WETZELS, 2006, p. 44):

MATCH-PROSODY (SI-PL), WEIGHT, ALIGN-FT-R, PRWD-R, TROCHEE >> BINARITY ( $\sigma$ )

MATCH-PROSODY (SI-PL), ALIGN-FT-R, PRWD-R >> WEIGHT

A maioria dos estudos fonológicos sobre o acento no PB se vale de teorias não lineares, com predomínio da Fonologia Métrica e da Fonologia Lexical. Tais estudos, contudo, encontram problemas teóricos e empíricos, muitas vezes decorrentes de restrições dos modelos adotados – e.g. o frequente emprego de explicações *ad hoc* e artifícios teóricos, como extrametricidade e marcação no léxico/representação subjacente, para tratar casos que não se enquadram na análise. Ainda permanece em aberto o debate sobre em que medida o acento é especificado no léxico, implementado por algoritmo ou sensível ao peso silábico. Este estudo pretende contribuir com a compreensão dos mecanismos de atribuição do acento, adotando uma perspectiva teórica inovadora com relação às apresentadas anteriormente.

15 *Sílabas pesadas* são aquelas em que a rima é complexa (ou seja, sílabas em que o núcleo é ocupado por ditongo ou vogal longa e/ou a coda é preenchida), contrapondo-se às *sílabas leves*, que apresentam rima simples (cf. HOGG; MCCULLY, 1987, p. 36-38).

16 O pé é um dos constituintes prosódicos na proposta de Nespor e Vogel (1986, p. 11), sendo superior à sílaba e inferior à palavra prosódica. Os pés podem ser classificados de acordo com sua estrutura interna. O pé trocaico, representado como (s w), é composto de sílaba forte (s) seguida de sílaba fraca (w); o pé jâmbico, ao contrário, é

### 3.5 Características acústicas e articulatórias do acento no PB

Fernandes (1977), avaliando as propriedades acústicas do acento no português na variedade paulista, identificou como relevantes a duração e em menor grau a frequência, seguida da intensidade. Ferreira Neto (2007), em estudo também sobre a variedade paulista, avaliou os seguintes parâmetros acústicos, procurando relacioná-los às funções distintiva e contrastiva do acento: intensidade, duração, frequência fundamental e qualidade vocálica (F1 e F2). O autor concluiu que a qualidade vocálica não bastaria como fator para a segmentação das palavras, como previsto pela função demarcativa do acento. Encontrou uma fraca correlação (1) entre frequência e intensidade e (2) entre acento, frequência e intensidade, separadamente e em conjunto. Por exclusão, concluiu que a duração seria o principal correlato acústico do acento. Ferreira Neto, contudo, não contemplou em sua análise as diferenças entre sílabas acentuadas e não acentuadas em cada um dos três padrões acentuais (oxítonas, paroxítonas, proparoxítonas), o que pode ter levado a uma redução no poder explicativo de seus resultados.

Major (1985) testou a produção do acento lexical por um falante baiano, um mineiro e um paranaense. Encontrou na duração o correlato mais confiável do acento. Alega que, em seus dados, intensidade e frequência apresentaram inconsistências e grande variabilidade:

In some utterances, mingograms revealed that the intensity of the pretonic was approximately equal to that of the posttonic, though the posttonic was shorter; or that the intensity and pitch of the pretonic were approximately equal to that of the tonic, though the pretonic was shorter; or that the intensity and pitch of the posttonic were approximately equal to that of the tonic, though the posttonic was shorter. Thus the relative pitch and intensity of the syllables may vary considerably, while the durational ratios remain fairly constant. This indicates that the primary correlate of stress in BP is length. (MAJOR, 1985, p. 261).

O autor descartou intensidade e frequência como correlatos do acento no português e, por essa razão, só apresentou os dados referentes à duração, reproduzidos abaixo, na FIG. 5.

---

composto de sílaba fraca seguida de sílaba forte, sendo representado como (w s). Nespor e Vogel (1986) adotam o modelo arbóreo de representação, com a notação (s) para sílaba forte e (w) para sílaba fraca. No modelo da grade métrica, utilizado, por exemplo, em Halle e Vergnaud (1987), adota-se a notação (\*) e (.), em lugar de (s) e (w). No presente trabalho, não adotaremos a noção de pé, que é controversa.

17 A extrametricidade é a mecanismo da Fonologia Métrica, proposto inicialmente por Liberman e Prince (1977), para lidar com dados que não se encaixavam perfeitamente nas regras. Funciona de modo a tornar as

	AVERAGE DURATIONS (msec)			DURATIONAL RATIOS			STANDARD DEVIATIONS (msec)			
	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	S <sub>2</sub> /S <sub>1</sub>	S <sub>2</sub> /S <sub>3</sub>	S <sub>1</sub> /S <sub>3</sub>	S <sub>1</sub>	S <sub>2</sub>	S <sub>3</sub>	
Speaker 1 (Bahia)	148	240	99	1.62	2.42	1.50	17	25	18	(N = 30)
Speaker 2 (Minas Gerais)	168	235	125	1.40	1.88	1.34	22	30	19	(N = 17)
Speaker 3 (Paraná)	164	255	111	1.55	2.30	1.48	20	26	17	(N = 38)

TABLE 1. Levels of significance: mean durations between all syllables (for each speaker) are significant to > 0.01.

FIGURA 5 - Média de duração e duração relativa em três variedades do português

FONTE: MAJOR, 1985, p. 261.

No dialeto carioca, Moraes (1998) relata que a manifestação do acento é complexa, uma vez que a combinação entre os parâmetros duração, frequência e intensidade é influenciada pelo contexto de ocorrência.

In Brazilian Portuguese, acoustic correlates of lexical stress vary as a function of a word's position in a utterance and, less significantly, as a function of the word's stress pattern. When a word is found in a weak position (i.e., inside a prosodic group), stress prominence is expressed as a combination of intensity and duration. (MORAES, 1998, p. 180).

No par de figuras que se segue, FIG. 6 e FIG. 7, reproduzidas de Moraes (1998), representam-se os correlatos acústicos do acento, respectivamente em início de sentença (posição fraca no domínio da entonação) e final de sentença (posição proeminente no domínio da entonação). As classes s1, s2 e s3 representam respectivamente a antepenúltima, a penúltima e a última sílabas da palavra, enquanto a única s0 (linha de F0 e coluna de proparoxítonas) se refere à necessidade de especificar a sílaba anterior à antepenúltima: como mostram os dados de Moraes, o acento lexical corresponde a uma queda em F0 após a produção da sílaba acentuada. Para demonstrar tal padrão, faz-se necessário, portanto, especificar F0 na sílaba anterior à acentuada, o que, nas proparoxítonas, corresponde à sílaba anterior à antepenúltima – s0. A relação entre acento e intensidade, por outro lado, consiste antes na diferenciação das postônicas por meio da redução de intensidade. Em duração, o mesmo acontece, mas de forma menos pronunciada: as postônicas são mais curtas que as pretônicas, e ambas são mais curtas que as acentuadas.

unidades classificadas como extramétricas invisíveis às regras de construção de constituintes (HALLE; VERGNAUD, 1987, p. 18).

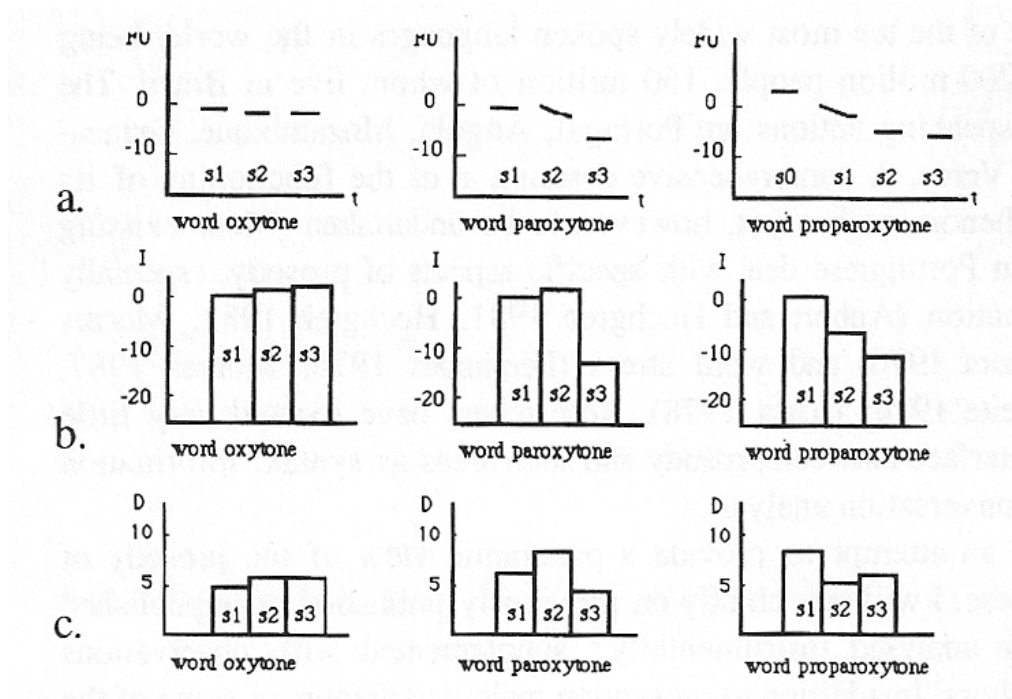


FIGURA 6 - Propriedades acústicas do acento lexical em início de sentença

a) F0 (quartos de tom); b) Intensidade (dB); c) Duração (cs)

FONTE: MORAES, 1998, p. 180

Comparando as FIG. 6 e 7, o autor destaca ainda que, em início de sentença (FIG. 6), praticamente não há diferença acústica entre as três sílabas nas oxítonas, ao passo que, em posição final de sentença (FIG. 7), a sílaba acentuada passa a ser sinalizada pela queda em F0 e aumento significativo em duração. Moraes infere que a diferença entre as duas situações seria devida à sobreposição da proeminência entonacional sobre a acentual, em fim de sentença. Comparando as FIG. 6 e 7, nota-se ainda um incremento na expressão acústica da proeminência: em posição final, os contrastes entre as sílabas acentuadas e não acentuadas são mais visíveis, principalmente em F0 e duração, intensidade em menor medida.

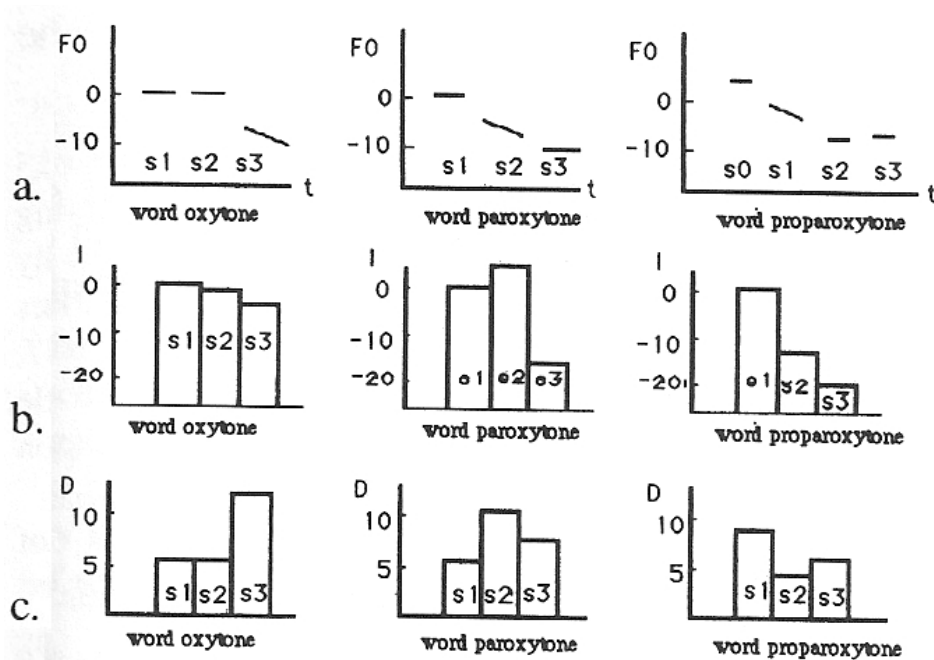


FIGURA 7 - Propriedades acústicas do acento lexical em final de sentença

a) F0 (quartos de tom); b) Intensidade (dB); c) Duração (cs)

FONTE: MORAES, 1998, p. 181

Infelizmente, o português brasileiro não conta, ainda, com descrições acústicas tão variadas e abrangentes como outras línguas, em especial com relação ao acento. A maioria dos estudos sobre o acento no português assume uma perspectiva fonológica, não se ocupando das características acústicas do fenômeno. De fato, estudos acústicos em Prosódia, no PB, são majoritariamente devotados à entonação e ao ritmo.

Ademais, os estudos fonéticos sobre o fenômeno acentual, como os que apresentamos anteriormente, em geral consideram apenas propriedades acústicas e não avaliam os mecanismos articulatorios envolvidos na acentuação. Não foram encontrados, até o momento, estudos sistemáticos sobre a articulação do acento no português no nível glótico. Tampouco foram encontrados estudos sobre a relação entre o acento no português e a movimentação da cabeça e da face. Mesmo em outras línguas, tais estudos ainda são escassos, imagina-se, pela maior dificuldade atrelada à obtenção de dados de movimento em comparação a dados acústicos.

Para alcançar o objetivo proposto de medição de propriedades acústicas e articulatorias envolvidas no acento, foi realizado, no presente estudo, um experimento de produção de fala. Avaliamos a variedade mineira. Tal experimento será relatado no cap. 6.

A análise acústica realizada neste estudo contemplará as medidas de duração, intensidade e frequência fundamental, que a literatura aponta como relevantes para a realização do acento no português. Dessa forma, poderemos promover uma comparação com os dados de Fernandes (1977), Ferreira Neto (2007), Major (1985) e Moraes (1998), mencionados na seção 3.5.

Pretendemos ainda analisar propriedades articulatórias relacionadas à produção do acento no nível glótico, por meio de medidas referentes à dinâmica de fechamento das cordas vocais, obtidas por eletroglotografia. Pretendia-se adicionalmente avaliar a relação entre o acento do português e a movimentação corporal superior, por meio de medidas do deslocamento da mandíbula, da cabeça e da dinâmica dos gestos faciais, coletadas pelo Optotrak, mas a análise de tais dados excederia os limites de tempo deste estudo.

O QUADRO 2, a seguir, retoma os principais aspectos levantados pelos estudos fonológicos e acústicos do acento no PB, apresentados nas seções 3.4 e 3.5. Os aspectos foram divididos em três níveis de análise – funcional, físico e cognitivo –, com fins estritamente metodológicos, que nortearão o desenvolvimento da análise multirrepresentacional que será aqui proposta.

<b>Nível de análise</b>	<b>Propriedades</b>
Físico	Correlatos acústicos e articulatórios
Cognitivo	Previsibilidade por regra, por peso silábico ou especificação no léxico
Funcional	Interação entre morfologia, fonologia e prosódia Produtividade de padrões Contraste, demarcação e hierarquização de unidades prosódicas (função distintiva, demarcativa, culminativa) Diferenças referentes ao sistema nominal e verbal

QUADRO 2 - Aspectos a serem abordados na análise multirrepresentacional do acento proposta neste estudo

Esses aspectos serão considerados sob uma perspectiva não baseada em traços e ou regras (ou qualquer mecanismo do tipo computacional, envolvendo unidades discretas e processamento serialista), mas antes em uma perspectiva multirrepresentacional, baseada no uso. O nível funcional será discutido no cap. 5. Os níveis físico e cognitivo serão discutidos também no cap. 5 e avaliados experimentalmente no capítulo 6.

## 4 Percepção

Este capítulo apresenta alguns aspectos da percepção da fala relevantes para o estudo em questão. Enfocamos, em especial, a complexidade envolvida na percepção da fala, a percepção de aspectos prosódicos e divergências teóricas quanto aos mecanismos perceptuais. Para desenvolver as ideias centrais deste capítulo, foram utilizados como fontes de referência Bregman (1994), Handel (1993) e Repp (1984), a quem remetemos o leitor para uma visão aprofundada sobre o assunto.

### 4.1 A complexidade da percepção

A produção da fala acontece a partir da integração de sons e de gestos. A percepção também é um fenômeno multimodal, pois se vale simultaneamente da audição e da visão. Entretanto, o principal sentido em que se baseia a percepção é, sem dúvida, a audição, um processo que, como veremos, é ativo e sujeito a influências contextuais.

Eventos sonoros são uma das formas básicas que utilizamos para interagir com o ambiente e são responsáveis pela compreensão da realidade espacial e temporal que nos cerca. Ao analisarmos o fenômeno da percepção sonora, nos deparamos com um fato intrigante: não há correspondência direta entre os eventos sonoros percebidos e sua configuração física. Na história do estudo científico da fala, especialmente após o surgimento de ferramentas de medição e visualização acústica, mostrou-se complexa a tarefa de isolar na onda acústica as unidades perceptuais supostamente nela instanciadas. Como um ouvinte consegue captar da fala de seu interlocutor as informações que este procurou expressar, tornando a linguagem humana uma estratégia de comunicação eficaz?

Uma das principais razões que tornam possível a comunicação é o fato de que a percepção é um processo ativo. Além dos sentidos, que captam do ambiente os estímulos sonoros e visuais, os processos cognitivos desempenham um papel fundamental na percepção da fala e são responsáveis pela decodificação de informações a partir do sinal da fala, notadamente um fenômeno complexo.

A onda acústica é inerentemente ambígua, em vista da continuidade e da sobreposição dos eventos sonoros. Como já mencionamos, não há correspondência simples e direta entre partes do sinal acústico e um segmento fonético. A fala envolve uma série de eventos acústicos que acontecem em menos de um segundo e que são unidos não sequencialmente, de forma complexa. Uma das principais fontes de complexidade na fala, responsável pela sobreposição de eventos sonoros, é a coarticulação, fenômeno decorrente da interação na movimentação dos articuladores e, conseqüentemente, entre partes sucessivas do sinal.

Apesar da complexidade do sinal da fala, podem ser identificadas regularidades tanto em termos de padrões de movimento e posicionamento dos articuladores (espaciais), quanto em termos da sequenciação dos segmentos no contínuo da fala (temporais).

Audition is a spatial sense as well as a temporal sense. [...] The ability to localize and thereby distinguish among sound sources is critical to our ability to perceive individual events. [...] time is fused with space to provide the framework for listening and that the rate of change temporally and spatially specifies the events themselves. (HANDEL, 1993, p. 174).

A percepção da fala é sensível a propriedades temporais, como taxa de articulação do enunciado, e fatores prosódicos, como a posição de um som na sentença. Os ouvintes depreendem os eventos do sinal acústico, baseando-se em propriedades da onda em um ponto do tempo e em relações entre partes da onda ao longo do tempo, a fim de gerar eventos perceptuais maximamente coerentes. É necessário, portanto, processar a onda acústica em duas dimensões temporais sobrepostas, que se referem à organização de padrões em curtos períodos de tempo e por longos períodos de tempo. Tal processamento deve ocorrer simultaneamente, uma vez que as propriedades acústicas de um mesmo evento sonoro e entre diferentes eventos sonoros não são independentes, mas antes se influenciam de forma mútua.

The reason that one single cue is not necessary and sufficient for identification is the interactive and context-dependent nature of sound production. The attributes of a sound are not independent. The vibration patterns, resonances, vibrato and so on are interactive; each aspect of sound production influences another. This influence spreads across a broad time span, and thus each aspect has an imprint on all others. For example, speaking rate and duration influence coarticulation and the harmonic spectrum. (HANDEL, 1993, p. 222)

A audição acontece em níveis diferentes de processamento cognitivo, por vezes utilizando todos ao mesmo tempo, por vezes, apenas um ou dois deles. Ela se baseia nas propriedades físicas dos sons, mas, em nível mais elaborado, se baseia também em propriedades dos sons que não guardam uma correspondência direta com a cadeia sonora (como brilho, *warmness*, etc.). Em um nível ainda mais avançado em termos de processamento, a audição de sons equivale à percepção de objetos e está atrelada à associação do evento com significados.

Se não é possível alcançar significados sem passar pelas propriedades acústicas, por outro lado, as propriedades acústicas só são percebidas dentro da estrutura oferecida pelo significado, em uma circularidade que é característica do processamento linguístico. A fala é percebida em termos do significado das cadeias sonoras, uma vez que a relação entre os sons da fala e seu significado é arbitrária.

O objetivo central de Handel (1993) é defender que os princípios subjacentes à percepção de eventos visuais são os mesmos que subjazem eventos auditivos. Retoma então os princípios de organização do campo visual propostos pela Gestalt, estendendo-os à modalidade sonora. Objetos tendem a ser agrupados por: (1) similaridade; (2) proximidade (espacial ou temporal); (3) continuidade (elementos que seguem na mesma direção); (4) movimentação conjunta; (5) simetria e fechamento (elementos que formam objetos simétricos e fechados). Essa proposta é compatível com os modelos teóricos adotados neste trabalho, uma vez que, se a percepção da fala e percepção de eventos visuais apresentam princípios de funcionamento semelhantes, reforça-se a ideia de que a linguagem se vale de processos cognitivos gerais.

#### 4.2 Percepção categórica

Ainda que, do ponto de vista da produção, a fala seja essencialmente contínua e apresente um alto grau de sobreposições, a percepção consegue isolar eventos e relacioná-los a conceitos bem delimitados, discretos. Tal tarefa envolve a segmentação do sinal da fala e a categorização de unidades significativas.

Categorization refers to processes in which different objects and events are placed or sorted into a small number of categories. Objects and events within a single category share a common property, although they differ in other irrelevant properties. [...] Categorizing is a ubiquitous perceptual cognitive process found in all domains and in no way specific to speech. (HANDEL, 1993, p. 266).

A categorização é um processo que intermedeia toda a experiência humana com o mundo e é fundamental à capacidade linguística e à tarefa de percepção da fala. “Perception of speech sounds requires categorization. This is because different instances of the same vocal gesture inevitably vary acoustically and, as a result, a wide range of physically different sounds must nevertheless represent the same phonemic entity.” (LIEBENTHAL et al., 2005, p. 1621).

Há certos eventos cuja percepção não se dá de forma gradual, mas em que, pelo contrário, ocorrem mudanças abruptas de uma classe para outra, ainda que os estímulos ambientais, em si, sejam contínuos. Esse tipo de percepção ficou conhecido como percepção categórica.

CP [categorical perception] is an important phenomenon in cognitive science because it involves the interplay between humans’ higher-level conceptual systems and their lower-level perceptual systems. [...] Another reason why CP is theoretically important is that [it] offers a potential account for how the apparently symbolic activity of high-level cognition can be grounded in perception and action. (GOLDSTONE; HENDRICKSON, 2009, [p. 1])

Considerando eventos sonoros dispostos em um contínuo, a percepção categórica consiste em estabelecer uma maior proximidade entre eventos pertencentes a uma mesma categoria e uma maior separação entre eventos pertencentes a categorias distintas. Em termos funcionais, trata-se de uma forma de interpretar a realidade quando é desnecessário processar detalhes finos, privilegiando o estabelecimento de grandes contrastes.

A percepção categórica pode ser testada por meio de experimentos de identificação, em que sujeitos são apresentados a estímulos interpolados entre dois sinais não ambíguos e devem classificá-los ou nomeá-los numa tarefa de escolha forçada. A partir do ajuste logístico dos resultados do teste, é traçada, para cada sujeito, uma curva psicométrica, que modela a probabilidade de resposta como função do parâmetro físico de interpolação dos estímulos. As curvas psicométricas consistem em gráficos que mostram a porcentagem de respostas conferidas a uma classe *versus* os estímulos apresentados aos sujeitos, como mostra

esquemáticamente a FIG. 8, a seguir. Tal curva reflete a dinâmica de categorização dos estímulos. No caso da percepção categórica, espera-se uma mudança brusca na identificação em função dos estímulos apresentados, como se pode ver sobretudo entre os estímulos 11 e 19 da FIG. 8.

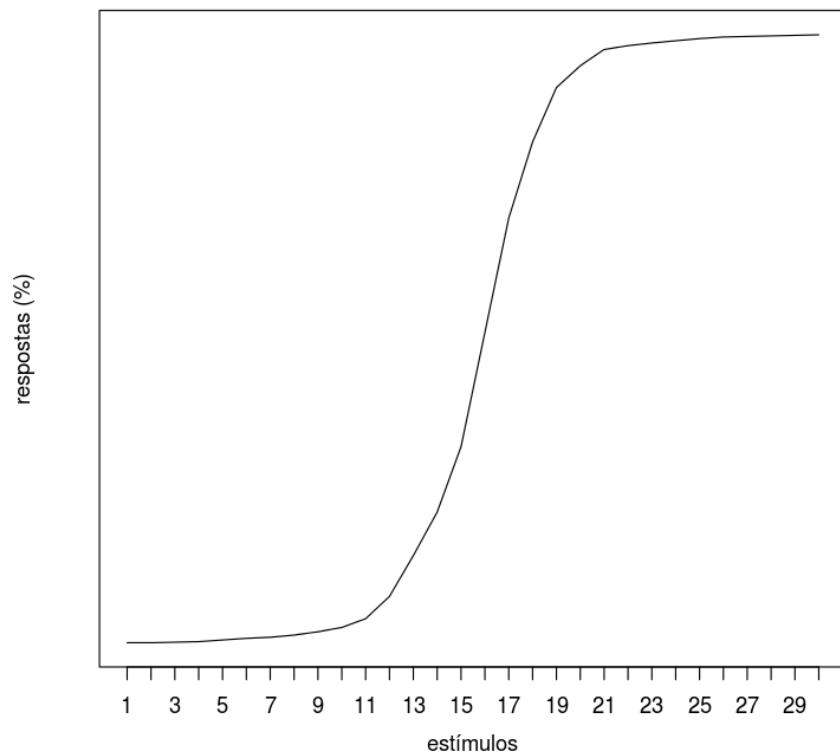


FIGURA 8 - Exemplo de curva psicométrica obtida em experimento de identificação

A curva psicométrica é definida unicamente pelos parâmetros seguintes:

- a) localização do intercepto ou limiar, região onde a curva psicométrica transformada atinge o valor de 50% no eixo vertical, que corresponderia a uma resposta ao acaso. O intercepto indica o ponto no contínuo dos estímulos onde ocorre a transição de percepção entre as categorias.
- b) a inclinação no ponto do intercepto da curva transformada. A inclinação se relaciona diretamente a quão “categórica” é a percepção ou quão abrupta é a transição entre as categorias.

Na percepção da fala, operamos tanto com a percepção categórica quanto com a percepção gradual, em que os eventos são percebidos de forma contínua. Além disso, é

notável que a percepção categórica não se restrinja à linguagem ou à dimensão sonora: pode acontecer na percepção de eventos sonoros e visuais linguísticos e não linguísticos.

Sabe-se que se, de um lado, a percepção categórica pode ser adquirida por meio de aprendizagem e, por outro, está presente no desenvolvimento humano desde os primeiros meses. Ao que tudo indica, consiste em um fenômeno em certa medida inato, em certa medida, adquirido. Sabe-se também que ela não é exclusiva à espécie: alguns animais exibem padrões de respostas a estímulos compatíveis com a percepção categórica observada em humanos. Ainda se conhece pouco sobre como a percepção categórica acontece em termos cerebrais.

#### 4.3 Percepção e aspectos prosódicos

Aspectos prosódicos exercem grande influência na percepção da fala. Uma das razões para isso é a natureza inerentemente temporal da audição. “The temporal nature of the auditory world lends a special importance to time and the timing of sounds. Time provides a framework for auditory events.” (HANDEL, 1993, p. 173). Eventos sonoros em geral são identificados com base na forma com que os sons se iniciam e terminam, na organização temporal de sons que se repetem e na regularidade do intervalo de ocorrência entre sons.

O ritmo tem grande importância na percepção da fala (cf. DOOLING, 1974), assim como a entonação (cf. VAISSIERE, 2005). Sabe-se também que os acentos lexicais ao longo de uma sentença funcionam como facilitadores no processamento de expressões linguísticas. Cutler (1976), a partir de testes de tempo de reação, conclui que o processo de compreensão de sentenças é influenciado pela previsibilidade dos acentos seguintes, por meio de sua relação com a estrutura melódica da língua.

Cutler (2005) retoma experimentos sobre a percepção do acento e aponta que as pistas acústicas mais consensualmente aceitas como relevantes são F0 e duração: uma maior duração é sempre associada à sílaba acentuada; ademais, a percepção do acento é influenciada pela trajetória de F0 e pelo ajuste temporal de F0 dentro da sílaba, uma vez que, em línguas com acento lexical, as sílabas acentuadas funcionam como potenciais âncoras para a proeminência no nível da sentença. Não há acordo quanto à influência da intensidade, que,

para alguns, seria relevante para a percepção do acento somente de forma indireta, como ênfase espectral: sílabas acentuadas apresentam mais energia nas regiões espectrais de maior frequência.

#### 4.4 Perspectivas teóricas

Até a década de 70, os estudos sobre a dimensão sonora da linguagem se voltaram prioritariamente para a produção da fala. A percepção era geralmente avaliada em pesquisas ligadas à Psicologia e Psicoacústica, áreas a que devemos grande parte dos conhecimentos básicos sobre processamento e armazenagem de estímulos em geral, entre eles, os sonoros. Naquele período, uma das razões para a separação dos estudos de percepção e produção da fala talvez envolvesse mais uma consequência teórica que uma divisão de saberes: o paradigma modularista e inatista adotado pela Gramática Gerativa tem como corolário a normalização dos estímulos na percepção da fala, que consiste na ideia de que a percepção linguística implica no mapeamento de eventos fonéticos variáveis em elementos abstratos invariáveis, na estrutura fonológica, em geral supondo a aplicação reversa das regras envolvidas na produção da fala. Nesse processo, seriam filtradas e descartadas as informações redundantes e presumidamente irrelevantes para o reconhecimento lexical, tanto linguísticas – como alofonias –, quanto extralinguísticas –, como características do locutor, do meio de comunicação e ambiente, etc. Com isso, surgiram propostas de explicar a percepção da fala a partir da produção.

A ideia de que padrões acústicos levam a inferências sobre a articulação ganhou corpo com o desenvolvimento da *Motor Theory of Speech Perception* (LIBERMAN et al., 1957), segundo a qual a forma com que articulamos padrões sonoros influencia a maneira com que percebemos esses mesmos padrões na fala de outras pessoas. Para essa teoria, a percepção categórica seria exclusiva à fala – como já mencionamos anteriormente, há fortes evidências de que não é – e sua existência se deveria à produção: no contínuo da fala, há pontos de descontinuidade, como as consoantes oclusivas, que são produzidas por manobras articulatorias distintas (envolvendo, por exemplo, lábios, no caso das labiais, e dorso da língua, no caso das velares).

Especialmente na década de 80, ganharam desenvolvimento modelos psicológicos de memória e categorização que propõem o armazenamento detalhado de informações (e.g. teoria dos múltiplos traços de memória) e modelos linguísticos que rompem com a concepção de um léxico formado por elementos abstratos e discretos (e.g. teorias baseadas no uso e em exemplares). A hipótese da normalização começou a ser questionada, com base em evidências experimentais de que a repetição exerce um papel inegável na memória linguística e de que detalhes considerados irrelevantes são armazenados pelos falantes. Detalhes fonéticos exercem influência em níveis superiores do processamento linguístico (HAWKINS, 2003).

Desde então, esforços vêm sendo feitos a fim desenvolver modelos que considerem que o armazenamento e processamento da linguagem envolvem redundância e detalhamento; que o uso da linguagem é determinante para a percepção e para a produção da fala.

## 5 Abordagem multirrepresentacional do acento

Neste capítulo, iremos analisar a proeminência acentual no português segundo princípios de teorias linguísticas multirrepresentacionais. Defende-se que tal análise apresenta vantagens sobre as abordagens anteriores, (1) por prescindir de mecanismos cognitivos específicos à linguagem, como módulos e regras; (2) por acomodar tanto os casos considerados regulares quanto os excepcionais; (3) por integrar as diversas dimensões linguísticas (morfologia, semântica, etc); (4) por incorporar a multimodalidade, contemplando características funcionais, acústicas e gestuais envolvidas na produção e percepção do acento. Antes de passarmos à análise do acento do português segundo modelos multirrepresentacionais, no entanto, iremos realizar um breve desvio cronológico, abordando os princípios de funcionamento do acento lexical no latim, precursor temporal do português. Defendemos que tal passo é essencial, uma vez que, em uma perspectiva dinâmica, o estado anterior do sistema – no caso, o latim – é determinante para sua configuração atual – o sistema acentual do português (GELDER; PORT, 1995). Dessa forma, a análise que será construída neste estudo se baseará nos dados do português brasileiro, mas irá considerar acessoriamente informações diacrônicas.

### 5.1 Percurso histórico do acento lexical português

O português, como língua românica, tem origens diretas no latim falado pelo povo romano. Chegaram até nós inúmeros registros de diferentes períodos e estilos, que permitem ao investigador explorar a trajetória de mudanças sonoras envolvidas nas sucessivas transformações da língua latina em português arcaico e, posteriormente, em português moderno. Neste tipo de estudo, preferência é dada a amostras textuais que deixam entrever indícios de como eram pronunciadas as unidades gráficas que nos restaram (e.g. *Appendix Probi* e obras normativas, *graffiti* e excertos de escrita do povo, composições literárias em verso). Inegável que este tipo de análise guarda uma parcela de especulação, mas, infelizmente, há muito já não havendo falantes de latim disponíveis, é a única possibilidade de vislumbrar aspectos sonoros dessa língua. No caso do tema em estudo, afirmações mais

seguras podem ser elaboradas, em vista da documentação abundante sobre o acento latino, que apresentava padrões de localização bem definidos. Apresentaremos, a seguir, uma breve descrição do acento latino e suas transformações ao longo do surgimento do português, com base em Faria (1970, 1995), Silva (1995), Teyssier (2001), Williams (1975) e Zágari (1988).

No latim, o sistema acentual era regular e obedecia à “regra das três sílabas”: palavras de três ou mais sílabas traziam acento na penúltima sílaba se esta fosse longa<sup>18</sup> (e.g. *senātus* “senado”) e, se fosse breve, na antepenúltima sílaba (e.g. *ānīma* “alma”); palavras dissilábicas sempre apresentavam acento na penúltima sílaba, independentemente de sua duração (e.g. *fīdes* “fé”, *lābor* “trabalho”). Havia, portanto, uma interação entre o acento e duração silábica, propriedade fonológica conhecida tradicionalmente como sensibilidade ao peso silábico.

Se o acento era previsível, regular, era também produtivo: palavras que viessem de outra língua – do grego, por exemplo –, se adaptadas à declinação latina, obedeceriam também a acentuação do latim, como conta Diomedes (*apud* FARIA, 1970).

Quanto ao aspecto funcional, o acento, no latim, seria demarcativo, pois serviria como pista para sinalizar a proximidade da fronteira da palavra em uma sílaba (paroxítonas) ou duas sílabas (proparoxítonas). As raras exceções à janela de localização do acento são resultado de processos erosivos de cancelamento final (e.g. *illic* < *illice*, *cupit* < *cupiuit*, como testemunha Prisciano *apud* FARIA, 1970). O acento no latim seria também culminativo: considerando o domínio de aplicação, sabemos que em cada palavra latina (exceto os clíticos) haveria uma sílaba mais proeminente que as demais. Sobre tal sílaba recaía o acento (*accentum*).<sup>19</sup> Por outro lado, o acento não era contrastivo, este papel sendo antes exercido pela duração silábica: e.g. *nītor* “brilho” x *nītor* “(eu) me apoio”; *lēuis* “liso” x *lěuis* “rápido”.

Em resumo, de um ponto de vista fonológico, o acento latino se caracterizava pela regularidade de localização, pela produtividade e pelo exercício de função demarcativa e culminativa, mas não contrastiva.

Em uma perspectiva de uso e de exemplares, como a que é adotada neste estudo, a estrutura gramatical surge a partir do mapeamento de experiências linguísticas armazenadas na mente dos falantes, pela construção de esquemas abstratos (BYBEE, 2010;

18 Nas palavras latinas, se relevante, a sílaba longa terá sua vogal marcada pelo mácron ( ¯ ) e a breve, pela braquia ( ˇ ).

19 “A sílaba que mais soa na palavra toda é a que tem acento” (“Illa syllaba plus sonat in toto uerbo, quae accentum habet.”, Pompeio *apud* FARIA, 1995, p. 35).

LANGACKER, 1987). Nessa perspectiva, a regularidade e a produtividade, como no caso do acento latino, emergiriam da forte relação circular entre aspectos das experiências linguísticas armazenadas – a proeminência acentual das palavras – e as abstrações construídas a partir delas – esquemas acentuais.

Note-se que, na perspectiva teórica adotada neste trabalho, as representações construídas pelos falantes a partir dos eventos linguísticos são dinâmicas. Dessa forma, os diagramas ilustrativos que serão apresentados a seguir devem ser interpretados como “retratos” do sistema. Na FIG. 9, a seguir, é mostrada uma forma de representação da emergência de esquemas abstratos, pela projeção ou abstração de uma rede de relações acentuais.

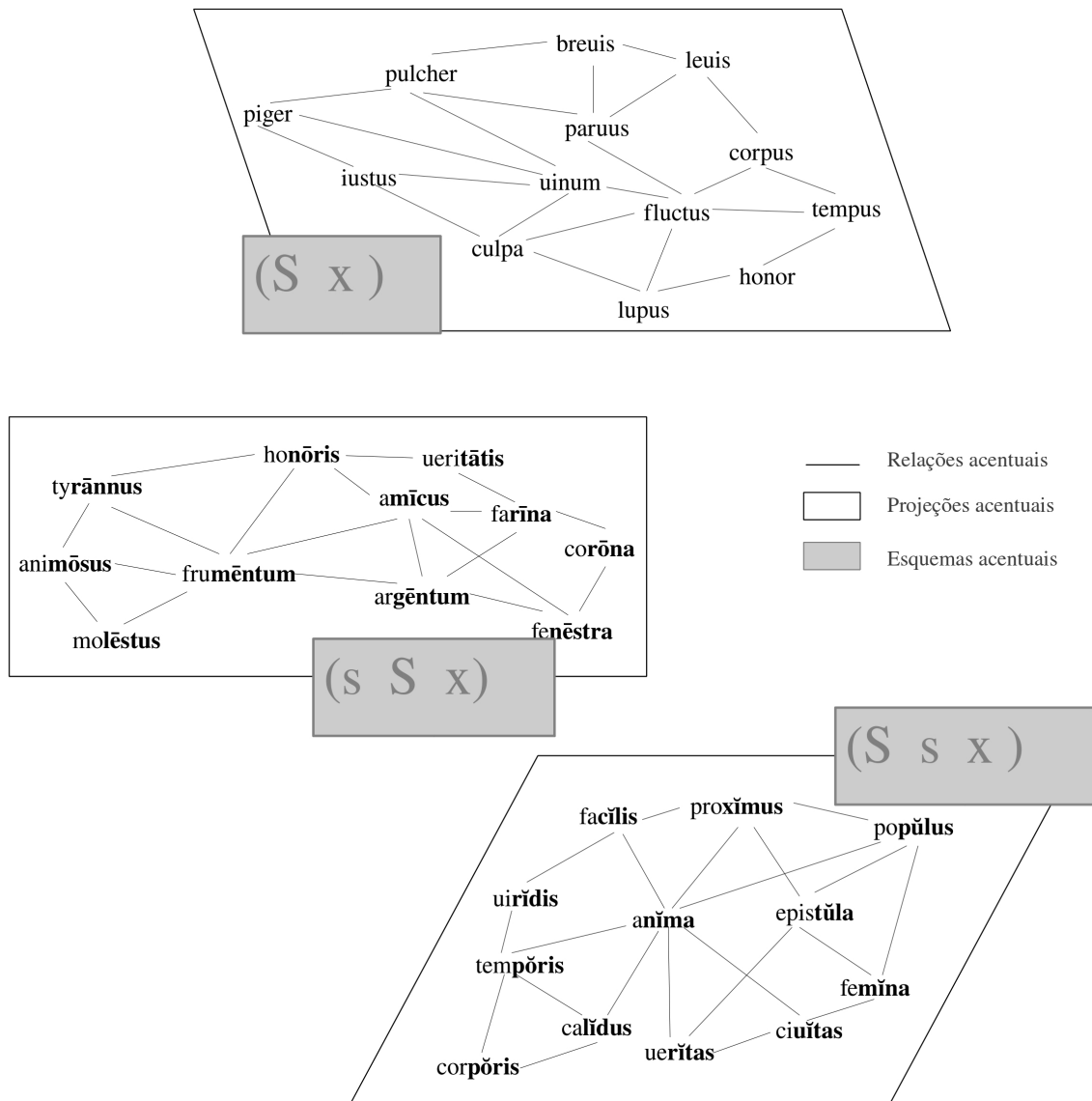


FIGURA 9 - Representação de redes e esquemas acentuais no latim

Esquemas acentuais, ilustrados por retângulos em cinza, são estruturas abstratas (*templates*) extraídas de padrões recorrentes na localização da proeminência em termos de contagem de sílabas e peso (i.e. quantidade) silábico. Ou seja, a partir de inúmeros elementos (como palavras e morfemas) que se comportam de forma semelhante (proeminência acentual na penúltima ou antepenúltima sílabas, a depender do peso da penúltima), emergem generalizações que refletem as tendências principais (de modo geral, [s S x] e [S s x]; em palavras dissílabas, [S x])<sup>20</sup>. Tais generalizações são projetadas abstratamente na forma de esquemas. A direção dos planos (paralelogramo com laterais inclinadas à esquerda, à direita ou verticais) reflete a coocorrência de um padrão acentual em um mesmo grupo de elementos, que leva à projeção abstrata das relações acentuais.

A análise do acento proposta neste estudo se fundamenta na ideia de que a organização acentual no português emerge a partir de relações semânticas, fonéticas e morfológicas consolidadas previamente, no latim. Argumentaremos que a coocorrência dos padrões regulares do sistema acentual latino com redes morfológicas robustas leva ao estabelecimento de conexões entre a localização do acento e a estrutura flexional das palavras no PB. Em um modelo de redes, tal mecanismo pode ser explicado como uma forma de *entrenchment*:

Entrenchment is straightforwardly identifiable as an adjustment in connection weights, brought about by the occurrence of a pattern of activation, which renders more likely the re-occurrence of the same or a comparable pattern. With respect to the system's movement through state space, entrenchment amounts to the emergence of an attractor.” (LANGACKER, 1999, p. 96).

A FIG. 10, a seguir, ilustra uma possibilidade de representação das relações morfológicas e acentuais latinas segundo um modelo de rede.

---

<sup>20</sup> Escolhemos arbitrariamente as letras “S”, “s” e “x” para representar, no sistema latino, respectivamente a sílaba que recebe proeminência, sílabas que não recebem proeminência e a sílaba final, que é ignorada na atribuição do acento.



Na FIG. 10, é proposta uma estrutura de redes formada a partir de conexões acentuais. Tal estrutura coexiste em meio a outras redes, como as estabelecidas pelos paradigmas de flexão, representadas na FIG. 10 como planos morfológicos. A relação estabelecida entre formas flexionadas (declinadas) de uma mesma palavra é mostrada na forma de uma rede local: um feixe de ligações tendo como centro uma forma de alta ocorrência. O pertencimento a uma mesma declinação é ilustrado pela inserção em um mesmo plano morfológico (e.g. *piger* e *pulcher* pertencem a um mesmo paradigma de declinação, enquanto *brevis* e *leuis* seguem outro paradigma). O compartilhamento parcial de uma declinação corresponde à sobreposição dos planos morfológicos (e.g. *piger* e *pulcher* pertencem a um paradigma de declinação parcialmente semelhante ao de *iustus*, *paruus* e *summus*). As relações que caracterizam um mesmo grupo ou classe gramatical são expressas pelas projeções morfológicas (e.g. *piger*, *pulcher* são adjetivos, enquanto *lupus*, *equus* e *agnus* são substantivos).

As representações propostas nas FIG. 9 e 10 contemplam apenas o sistema nominal, por uma restrição de espaço, e se aplicariam também ao sistema verbal, mas com um maior número de conexões entre os elementos, devido ao maior nível de complexidade flexional na morfologia verbal.

Como dissemos acima, a respeito da FIG. 9, esquemas acentuais abstratos são construídos a partir de padrões recorrentes na localização da proeminência acentual em termos de contagem de sílabas e peso silábico. O peso silábico, dessa forma, funcionaria como um atrator acentual no sistema latino. O *entrenchment* dos esquemas acentuais (FIG. 9) com a rede morfológica densa e coesa do latim (FIG. 10) resultaria em redundância nas conexões, com o surgimento de novo atrator no sistema acentual – ou seja, passariam a coexistir dois atratores, peso silábico e rede morfológica. Dessa forma, explica-se porque a perda da quantidade vocálica, que tinha impacto direto no peso silábico, não desestabilizou o sistema acentual na passagem para o português, como veremos a seguir.

A natureza fonética do acento no latim é muito debatida e talvez não seja possível determiná-la com segurança, em vista da ausência irremediável de material fônico daquela época. Contudo, a maioria dos estudiosos sugere que o acento era realizado por uma mistura das propriedades intensidade e altura, ainda que uma acessória à outra.

Pouco se conhece sobre a entoação latina, mas há indícios de que o ritmo da língua, desde o nível da palavra, até o nível da sentença, era fortemente baseado na quantidade

(FARIA, 1970). Alguns indícios sobre a entonação e acento podem ser inferidos da metrificação de obras em verso, em que cada pé era dotado de um *ictus*<sup>21</sup>. Este seria uma proeminência de nível superior ao acento, com o qual nem sempre se fazia coincidir. Como vemos nos versos a seguir (*Eneida*, IV, 439-443), o *ictus* de cada pé (sílabas sublinhadas) não segue a acentuação das palavras (negrito), mas a duração silábica, recaindo sempre sobre sílabas longas:

**nōx** rūīt| Āenē|ā nōs |**flē**ndō |dūcīmūs| **hō**rās  
**hīc** lōcūs| **ē**st pār|tīs ūbī| **sē** vīā| **fin**dīt ĩn| āmbās  
dēxterā| **quāe** Dī|tīs **māg**|nī sūb| mōenīā| **tē**ndīt  
**hāc** ĩter| Ēlysī|ūm nō|bīs āt| **lā**evā mǎ|lōrūm  
ēxēr|cēt **pōe**|nās **ēt** ād| īmpīā| Tārtārā| **mī**ttīt

É conhecida, contudo, uma tendência a fazer recair o *ictus* sobre o acento lexical nos dois últimos pés do verso, especialmente no penúltimo, o que poderia ser uma indicação de estrutura entonacional, de nível ainda superior ao *ictus*. Nesse sentido, o acento possivelmente seria usado como hierarquizador (função culminativa), interagindo com a quantidade vocálica e com o *ictus* para ancoragem da estrutura entonacional sobre a estrutura segmental. A título de ilustração, retomamos a seguir o primeiro verso do poema supracitado, acrescentando uma proposta de representação fonológica métrica, na FIG. 11, abaixo.

<sup>21</sup> Proeminência de cada pé métrico no verso, recaía sempre sobre a primeira sílaba do pé. Não há consenso sobre a natureza fonética do *ictus* no verso latino.

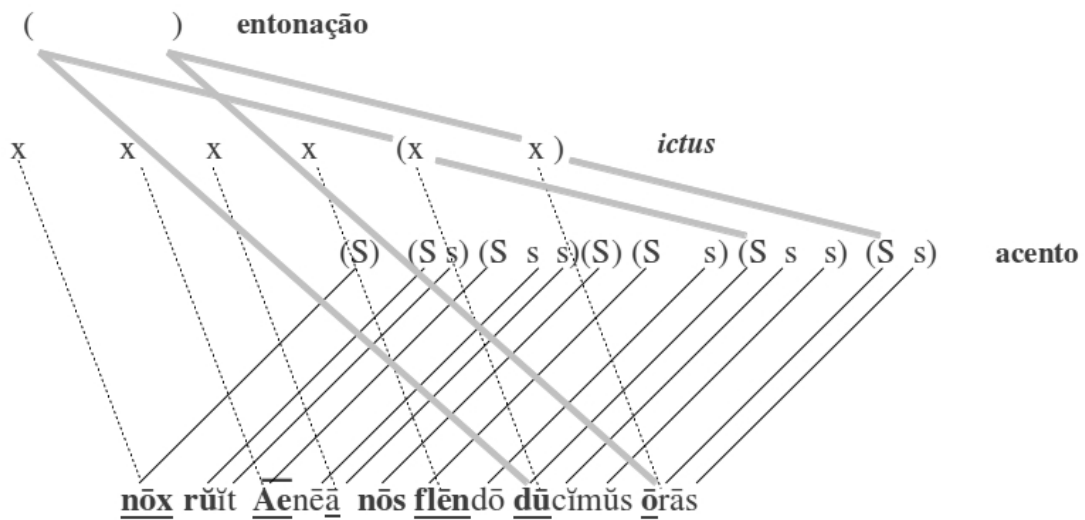


FIGURA 11 - Níveis prosódicos na metrificação latina

Um dos fenômenos sonoros mais marcantes da dissolução da língua latina, do qual as consequências se fizeram sentir em todas as línguas românicas, foi a perda da quantidade vocálica como propriedade distintiva – i.e. sua desfonologização. No caso da passagem do latim para o português, o percurso ocasionou a mudança de um sistema de dez vogais em contraste (“ī, ē, ā, ō, ū, ī, ē, ā, ō, ū”) para apenas sete (“i e ε a o o u”, em posição acentuada, considerando o português brasileiro atual). Tal mudança implicou, não só na redução do paradigma vocálico, por meio da fusão de categorias, mas também em alterações na qualidade das vogais médias, que passaram a distinguir-se, aos pares, por níveis mais sutis de altura na articulação (e x ε ; o x o). A consequência mais importante desse fenômeno para o raciocínio aqui desenvolvido diz respeito ao que interpretamos como uma desestabilização do sistema de acento regular latino. Como a quantidade vocálica tinha influência na quantidade silábica<sup>22</sup> e esta guardava forte correlação com a localização do acento na palavra, a desfonologização da quantidade vocálica não pode deixar de ocasionar uma mudança na funcionalidade do acento, que passou a desempenhar o papel contrastivo – mas a regularidade e previsibilidade de sua localização se perdeu na história, preservada apenas como fóssil nas próprias palavras. Nessa perspectiva, não haveria um mecanismo único responsável pela localização do acento no

<sup>22</sup> Não havia uma relação unívoca entre quantidade da vogal e da sílaba, visto que esta última era também influenciada por sons não nucleares da sílaba: e.g. em *mens* “mente”, independentemente da duração da vogal “e”, a sílaba seria longa, devido ao encontro de duas consoantes, “ns”, seguindo seu núcleo.

português. As aparentes relações entre estrutura silábica e localização do acento podem ser interpretadas como resquício dos *templates* acentuais do sistema latino. No português, a localização do acento passou a ser definida também com referência a palavras particulares, podendo mesmo ser propriedade estendida a unidades menores que a palavra, como morfemas<sup>23</sup>, que emergem na gramática a partir de relações semânticas e fonéticas. A estrutura morfológica passou a interagir diretamente com a acentuação.

Se há quem defenda que o acento latino fosse realizado por mais de um parâmetro acústico, fato é que a perda da quantidade vocálica libertou a propriedade de duração de seu papel contrastivo anterior e esta passou a se correlacionar com a realização do próprio acento. Tal relação já existia, uma vez que as sílabas acentuadas nas paroxítonas eram sempre longas, mas inverteu-se o jogo entre a propriedade principal e a acessória na relação acento-quantidade. Como vimos no cap. 2, estudos do acento no português brasileiro apontam como correlatos fonéticos principais justamente duração e intensidade, o que seria condizente com a hipótese descrita nos parágrafos anteriores.

Ao longo da transformação do sistema acentual latino no que veio a se tornar o sistema acentual português, o surgimento e fortalecimento do padrão oxítono ocorreram principalmente como consequência de processos fonológicos redutivos (e.g. apagamento de sílaba átona final ou medial, cf. Teyssier (2001)) e empréstimos de outras línguas (francês, árabe e línguas indígenas e africanas, no português brasileiro).

No surgimento do português, mesmo com a desfonologização da quantidade, motivação primária do acento latino, foi preservada nas palavras e morfemas a localização acentual. Algumas exceções dizem respeito a mudanças no padrão acentual não relacionadas a empréstimos ou a cancelamentos finais, que atribuiremos a efeitos das conexões em rede sobre a estrutura gramatical. Por exemplo, nas formas latinas de imperfeito na primeira e segunda pessoas do plural (e.g. *amabāmus* “amávamos” e *amabātis* “amáveis”), o acento caía sobre a flexão de tempo (-bā-), deslocando-se para a sílaba anterior no português. A explicação desse fenômeno em modelos fonológicos tradicionais é intrincada – e.g. Lee (2007) recorre ao princípio de uniformidade paradigmática, que busca explicar casos em que palavras se comportam diferentemente do esperado, seguindo o paradigma morfológico ao invés de se sujeitarem a regras fonológicas.

---

23 Para uma abordagem do acento do português como propriedade dos morfemas, cf. Cagliari (1999).

Em um modelo de exemplares, tal fenômeno se manifesta naturalmente como consequência de efeitos de frequência lexical de tipo, em um sistema morfológico concebido como estrutura emergente a partir de relações em rede (BYBEE, 1995). Se a estrutura linguística emerge da armazenagem e conexão das ocorrências de uso, a frequência de uso e de tipo teriam impacto na organização gramatical, sendo determinantes, respectivamente, para a força lexical e para a produtividade de padrões (BYBEE, 1985). Segue daí que os padrões de menor frequência de ocorrência, como as formas verbais do imperfeito de primeira e segunda pessoas do plural, se mostram mais suscetíveis ao nivelamento analógico (BYBEE, 2001), passando a integrar o paradigma regular, em que o acento recai sobre a última sílaba da base, como no caso apresentado, em já não havendo restrições quanto à localização do acento.

## 5.2 Abordagem multirrepresentacional

Hyman (1977) propõe, na forma de uma tipologia funcional, a existência de dois tipos de sistema acentual: o primeiro, baseado no acento lexical – cuja localização na palavra seria livre e imprevisível, como na língua *asmat*, da Nova Guiné –, derivaria historicamente do segundo, baseado no acento gramatical – com localização predizível, seja demarcando a fronteira das palavras (função demarcativa, como no polonês), seja as fronteiras de morfemas (função morfológica, como no russo). Analisando diversas línguas não aparentadas, Hyman (1977) sugere que o percurso histórico de transformação do acento gramatical em lexical geralmente resultaria de fenômenos como a perda do peso silábico, o cancelamento de vogais postônicas e a epêntese vocálica. Os três fenômenos mencionados por Hyman aconteceram com o surgimento do português – os dois últimos ainda acontecem – e, nessa perspectiva, interagiriam com os padrões de localização do acento na língua. Hyman reconhece a possibilidade de co-exercício de mais de uma função pelo acento no mesmo sistema acentual, o que parece se adequar ao caso do espanhol, mencionado pelo autor, e do português, caso do presente estudo. Procuraremos mostrar, a seguir, que o acento no PB tem função demarcativa, interage com a estrutura morfológica, mas não é estritamente determinado por esses dois fatores e desempenha um papel na estrutura lexical da língua. Exerceria, portanto, tanto a função lexical quanto gramatical, possibilidade que Hyman reconhece e que poderia apontar

para a necessidade de inserir um terceiro novo tipo de sistema acentual em sua tipologia.

O acento, no português, funciona como delimitador de um núcleo lexical, ancorado no plano silábico, e estabelece uma conexão entre as estruturas segmental, entonacional e conceitual. Como consequência do papel de delimitador lexical exercido pelo acento, verifica-se no PB uma tendência à preservação da sílaba acentuada em fenômenos de variação e mudança linguística. Nessa língua, fenômenos sonoros afetam preferencialmente as sílabas átonas, conservando o núcleo lexical, ancorado na sílaba acentuada. Ademais, deslocamentos acentuais são raros (e.g. poucos casos como *xerox* x *xerox*; *ruim* x *ruim*). Línguas em que o acento “flutua”, dependendo da construção sintática, apresentariam outro tipo de relações e outra funcionalidade por trás do fenômeno acentual. Considerando a perspectiva teórica proposta pela Linguística Cognitiva, a acentuação no PB pode ser interpretada em termos de uma relação cognitiva de figura (*figure*) e fundo (*ground*).

[...] figure/ground organization is a valid and fundamental feature of cognitive functioning. By the assumptions of cognitive grammar, the prevalence of figure/ground organization in conceptual structure entails its importance for semantic and grammatical structure as well. [...] phenomena generally recognized as instances of figure/ground alignment might be regarded as special cases of the more general and pervasive process of comparison. It would then be possible to characterize a figure as the target of scanning at some level of organization. (LANGACKER, 1987, p. 120-121).

A sílaba acentuada, à qual é conferida maior proeminência sonora, se destacaria também pela maior saliência cognitiva – i.e. seria a figura. Já as sílabas não acentuadas, menos proeminentes em termos sonoros, seriam menos salientes cognitivamente e desempenhariam o papel de fundo na comparação estabelecida pelo acento. Nesse contexto, a preservação do acento pelos fenômenos sonoros pode ser interpretada como uma estratégia de preservação da saliência cognitiva para otimização do acesso lexical.

O acento estabelece uma conexão entre as estruturas segmental, entonacional e conceitual. Resultados de Moraes (1998) apontam para um aumento dos contrastes em F0, intensidade e duração entre sílabas átonas pretônicas, sílabas acentuadas e sílabas átonas postônicas.

A conexão entre as estruturas segmental, entonacional e conceitual, intermediada pelo acento, se dá, no PB, por meio da morfologia, podendo ser estabelecida uma relação entre a estrutura morfológica e a acentuação: tanto a derivação por sufixação quanto a flexão geralmente implicam no acréscimo de um morfema à margem direita da base, que, como

vimos, corresponde à janela de localização do acento. Iremos apresentar argumentos de que os dois processos morfológicos têm consequências diferentes para a acentuação no português.

A relação entre derivação e acentuação diz respeito essencialmente ao papel dos sufixos na determinação da localização do acento, o que pode se dar de formas variadas. Há sufixos que sempre recebem o acento (e.g. *-ção*, *-ismo*), outros que não recebem o acento, mas afetam a acentuação da base (e.g. *-ico*, *-imo*), e outros ainda cuja adição não altera a acentuação da base (e.g. *-o*), como se pode ver nos exemplos abaixo:

*civil* → *cívico*; *civismo*

*cantar* → *canto*

Por outro lado, prefixos em geral não influenciam a acentuação primária da palavra derivada, mesmo em monossílabos (e.g. *tripé*, *enfim*), o que estaria relacionado ao fato de que a localização do acento no português considera a janela das últimas sílabas, na margem direita da palavra. Palavras não derivadas se apresentam com o acento próprio da base.

A flexão influencia a acentuação de forma aparentemente semelhante à derivação. Os morfemas flexionais, como os sufixos, são adicionados à margem direita da base e podem determinar a localização do acento. Alguns morfemas flexionais atraem para si o acento (e.g. *-ia-*, pretérito imperfeito, segunda e terceira conjugações), outros alteram a acentuação na base (e.g. *-mos*, da primeira pessoa plural) e outros não interferem na acentuação (e.g. *-o*, da primeira pessoa singular, presente).

Dos dados apresentados, observa-se que os morfemas adicionados à margem direita da base, sejam eles derivacionais ou flexionais, podem determinar a acentuação da palavra, o que acontece geralmente quando tais morfemas consistem em uma mais sílabas e explicaria por que a maior parte das regularidades na localização do acento no português está relacionada a padrões morfológicos (cf. CAGLIARI, 1999).

As abordagens tradicionais consideram as generalizações possíveis sobre o acento no PB como um reflexo de regras ou restrições e, portanto, como causa determinante da regularidade verificada nesse sistema acentual (cf. estudos reunidos em ARAÚJO, 2007). As abordagens de uso, ao contrário, buscam interpretar as generalizações como resultado da rotinização de padrões. Os modelos teóricos multirrepresentacionais adotados neste estudo assumem que padrões segmentais e prosódicos interagem com a morfologia em um modelo de rede de modo a configurar a gramática prosódica (BYBEE, 1985, 2010). Em uma análise multirrepresentacional, defende-se aqui que padrões gerais com relação ao acento no PB são

resultado de processos de categorização em vários níveis da organização gramatical. Tal categorização acontece por meio de redes de conexões entre exemplares, formadas e fortalecidas a partir da recorrência na correspondência de padrões acentuais e morfológicos.

Uma análise semelhante à proposta neste estudo foi delineada por Farrell (1990), que se vale do modelo cognitivo de Langacker para analisar o acento em não verbos do espanhol. Em consonância com Farrell, sugere-se, aqui, que o acento lexical integra esquemas de nível alto de abstração e complexidade, como morfologia, fonotática e sintaxe, sendo parte da informação transmitida por ocorrências de uso. Tanto os morfemas derivacionais quanto os flexionais contribuiriam para o estabelecimento de redes de conexões. No entanto, as consequências funcionais da derivação e da flexão na dinâmica acentual do PB são diferentes. A morfologia flexional, mas não a derivacional, proporciona o estabelecimento de redes acentuais em exemplares com maior grau de coesão, pois integram a base e o morfema flexional. O fato crucial para a diferença entre flexão e derivação consiste no caráter idiossincrático da última e na sistematicidade da primeira. Na derivação, são construídas redes entre os sufixos semelhantes e entre as bases às quais eles se ajuntam, mas não entre bases e sufixos simultaneamente, uma vez que o domínio de ocorrência dos sufixos é restrito. Na flexão, por outro lado, as redes formadas são mais complexas e mais fortes, devido à sistematicidade com que as flexões ocorrem no paradigma verbal.

As FIG. 12 e 13, a seguir, seriam uma proposta de representação integrada das redes de conexão morfológica, semântica e prosódica. Note-se que, nessas figuras, as sílabas foram classificadas como fortes (s) ou fracas (w) considerando somente sua proeminência relativa dentro da palavra (i.e. se acentuada ou átona), não a formação de pés tradicionalmente proposta em estudos anteriores sobre o acento do PB. A FIG. 12 representa uma rede de conexões morfológicas e acentuais no sistema verbal.

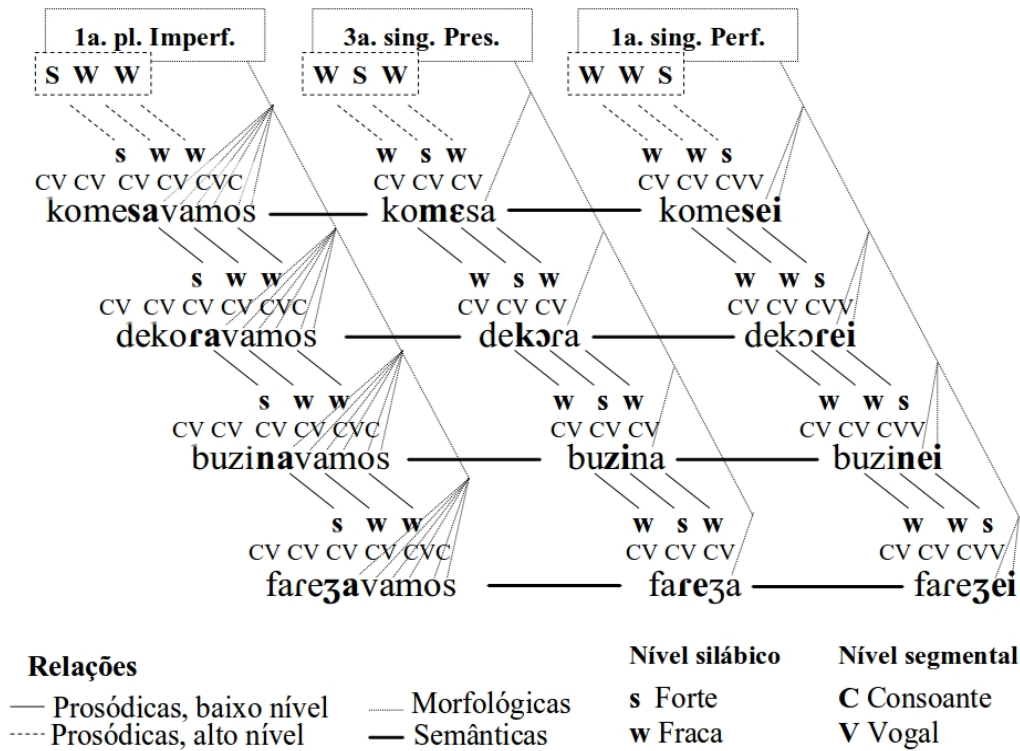


FIGURA 12 - Estrutura prosódica emergente e esquemas de nível alto para proparoxítonas, paroxítonas e oxítonas em verbos

A FIG. 13, por sua vez, mostra uma rede de conexões morfológicas e acentuais no sistema nominal.

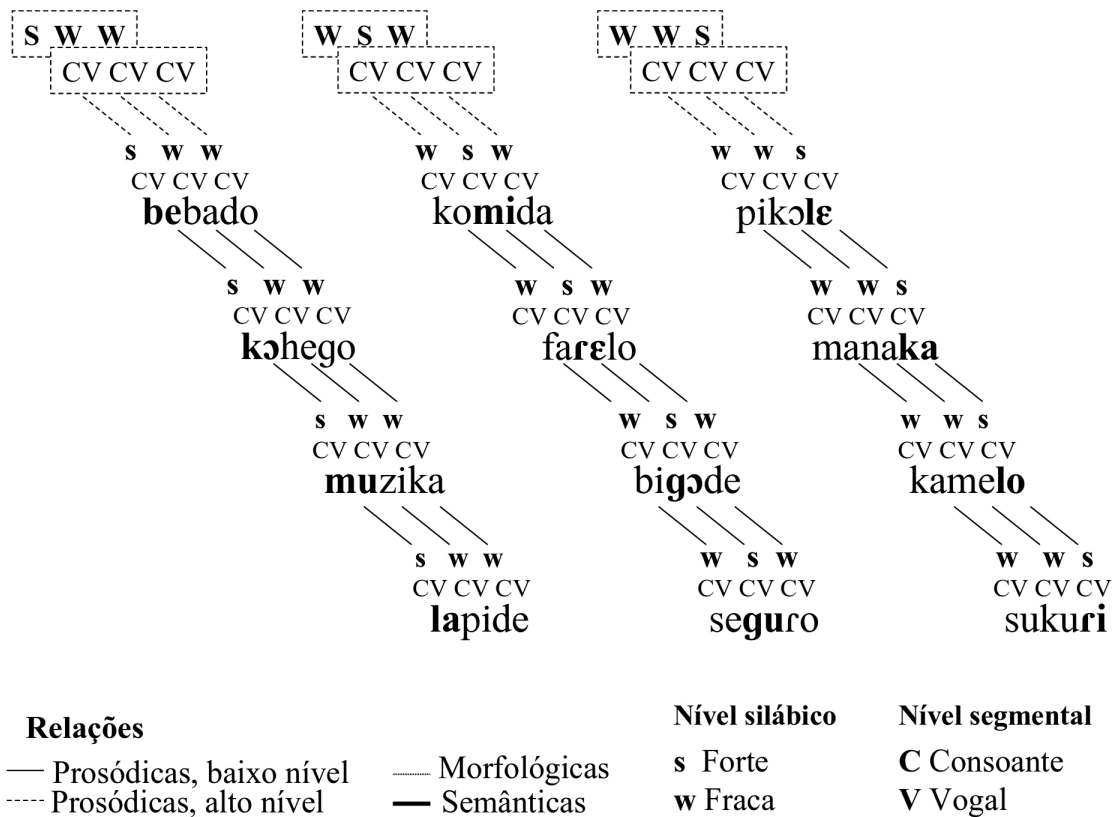


FIGURA 13 - Estrutura prosódica emergente e esquemas de nível alto para proparoxítonas, paroxítonas e oxítonas em não verbos

As FIG. 12 e 13 representam esquemas gramaticais, com destaque para esquemas prosódicos de maior e menor nível de abstração, que emergem a partir de relações em rede estabelecidas entre exemplares, respectivamente para verbos e não verbos. A partir do uso efetivo dos itens lexicais, são promovidas generalizações de naturezas diversas – dentre as quais generalizações relativas à estrutura silábica da língua, à estrutura morfológica, à localização do acento – e que correspondem ao conhecimento gramatical. Essas generalizações que emergem do uso linguístico levariam à construção de esquemas de nível baixo a alto, referentes à fonotática (i.e. organização da estrutura silábica), à morfologia e ao padrão acentual da língua específica.

Sugere-se, nesta análise, que, no PB, tais esquemas se integrariam no sistema verbal, originando um esquema abstrato e complexo, determinante da atribuição do acento. O mesmo não ocorreria no sistema não verbal, em que esquemas morfológicos flexionais não correm em paralelo aos esquemas acentuais. Assim, defende-se neste estudo que as particularidades da acentuação em formas verbais frente às não verbais podem ser atribuídas a diferenças no

mecanismo de armazenagem e processamento dos itens lexicais (BYBEE, 1985, 1995).

No PB, os verbos apresentam uma rede de informações morfológicas complexa e densa, que surge a partir da flexão verbal, responsável por expressar simultaneamente as noções de voz, aspecto, tempo, modo e concordância de número e pessoa, com certo grau de fusão. Comparativamente, relações flexionais nos não verbos são reduzidas e mais fracas, expressando apenas a noção de número (plural) e, algumas vezes, de gênero. A flexão de plural nos nomes acontece, em sua forma mais básica, pela presença de um morfema <-s> no fim das palavras – havendo formas mais complexas de plural no PB, mas que não vêm ao caso para os presentes propósitos. A presença do morfema de plural <-s> não alteraria a localização do acento. Mesmo nos casos em que se verifica a epêntese vocálica antes do morfema de plural, para evitar encontros consonantais indesejados (e.g. *amor*, *amores*, evitando o encontro *-rs*), a nova sílaba criada ao fim da palavra não afeta o padrão acentual. Nos nomes, portanto, haveria um acento *default* para a forma não flexionada, alterado apenas com a derivação, mas não com a flexão. Na derivação nominal, a localização do acento seria determinada pelos sufixos: o acento seria uma propriedade gramatical, morfológica.

Ao contrário dos não verbos, os verbos não apresentam um acento *default*, pois devem sempre ocorrer acompanhados de flexões que irão determinar o acento. Ademais, nos verbos, a vogal epentética entre duas consoantes de sílabas distintas pode opcionalmente receber o acento – (eu) **op**[i]to x **op**[i]to; (eu) **indig**[i]no ou **indig**[i]no –, havendo deslocamento do acento em consequência da ressilabificação, como testado por Cantoni e Cristóforo (2008). Tal fenômeno apresenta um desafio para modelos que assumem o processamento serial de regras fonológicas. Por exemplo, na análise de Lee (1992), o acento é definido como uma regra lexical (ou seja, que se aplica ainda no léxico, antes da passagem pelo componente sintático), enquanto a epêntese seria uma regra pós-lexical, aplicando-se somente depois de o *input* ter passado pelas regras lexicais e pelo componente sintático. Na Fonologia Lexical, em que se baseia a análise do autor, uma regra lexical não pode se aplicar depois de uma regra pós lexical, surgindo, assim, um problema de ordem teórica ao lidar com o fenômeno mencionado acima.

A maioria das conexões lexicais entre não verbos é de natureza derivacional. Pelo acréscimo de sufixos, itens são formados a partir de outros, com os quais passam a manter relação no léxico. Contudo, essa relação seria mais fraca que a flexional, possivelmente porque a representação mental do novo item formado por derivação dispõe de maior

autonomia conceitual que as formas flexionadas e, portanto, afasta-se mais da representação do item de que foi originada. Em vista disso, pode-se sugerir que redes complexas e densas, geradas principalmente pela flexão, seriam responsáveis por um maior grau de integração em esquemas de nível alto, em comparação com as redes da derivação. Com base nessa hipótese, é possível demonstrar como os esquemas acentuais que surgem a partir de formas verbais no PB tendem a ser anexados aos esquemas morfológicos que lhes são paralelos, enquanto que, nos não verbos, o mesmo não se aplica, pois não há paralelismo entre esquemas morfológicos e prosódicos. Portanto, o padrão acentual distinto das formas verbais se deve a suas redes flexionais mais fortes, que interagem com a prosódia em esquemas de nível alto.

A Fonologia de Uso (BYBEE, 1995, 2001) assume que as palavras são as unidades básicas de armazenamento e que tais unidades representacionais são ricas em redundância e detalhes. A análise a ser desenvolvida neste estudo sugere que informações sobre o acento integram a representação mental. Defende-se, portanto, que o acento no PB é lexicalmente especificado, o que corrobora a ideia de que informações redundantes estão presentes na representação mental.

Argumentamos que, se a informação sobre o acento está presente na representação lexical, é possível alcançar uma análise compreensiva da produtividade de padrões acentuais, tanto na morfologia verbal, quanto na não verbal. Experimentos em percepção de fala testando a natureza lexical do acento foram elaborados de forma a contribuir com essa análise e serão apresentados no cap. 6.

A produtividade de padrões, para a Fonologia de Uso, é diretamente influenciada pela frequência de tipo (BYBEE; HOPPER, 2001). Para avaliar a frequência de tipo dos padrões acentuais no PB, foi realizada uma coleta de dados preliminar junto ao banco de dados do Projeto ASPA (Avaliação Sonora do Português Atual)<sup>24</sup>, um *corpus* de padrões sonoros do PB. A TAB. 1 a seguir mostra a frequência de tipo obtida nessa *corpus* para cada uma das três posições do acento (proparoxítonas, paroxítonas e oxítonas), em verbos e não verbos. Os monossílabos foram excluídos dos cálculos, pois neles não se estabelece relação de proeminência acentual no nível da palavra, que é nosso principal interesse neste estudo.

---

24 <<http://www.projetoaspa.org>>.

TABELA 1 - Frequência de tipo de cada padrão acentual em verbos e não verbos

Posição do Acento	Verbos		Não verbos	
	N	%	N	%
Proparoxítonas	482	1,2	11.389	14,6
Paroxítonas	27.730	71,4	52.271	67,1
Oxítonas	10.617	27,3	14.200	18,2
<b>Total</b>	<b>38.829</b>	<b>100,0</b>	<b>77.860</b>	<b>100,0</b>

Considerando os dados dispostos na TAB. 1, verifica-se que a distribuição do número de palavras não é igual para os três padrões acentuais, tanto no grupo de itens que são somente verbos ( $\chi^2[2] = 29.308$ ,  $p < 0,001$ ), quanto no grupo de itens que não são verbos ( $\chi^2[2] = 40.182$ ,  $p < 0,001$ ). As palavras paroxítonas apresentam frequência de tipo significativamente maior que as demais e correspondem ao padrão acentual mais produtivo, como é frequentemente afirmado nos estudos sobre o acento no PB mencionados na seção 2.2. Esse fato é corroborado por uma análise da localização do acento em neologismos (cf. ALVES, 1994), em sua maioria paroxítonos.

Comparando os grupos de verbos e não verbos, vemos que há uma diferença significativa na distribuição de oxítonas, paroxítonas e proparoxítonas entre itens que são verbos e não são verbos ( $\chi^2[2] = 5.642$ ,  $p < 0,001$ ), o que seria um indício de que o sistema verbal e o sistema não verbal apresentam dinâmicas acentuais distintas.

Para avaliar a frequência de ocorrência, usamos a estratégia de separar os itens em alta frequência e baixa frequência, considerando como critério de corte o valor de frequência de 8007 ocorrências (o que equivale ao critério de 35 por milhão, adotado por Bybee (2002b), ao ser adequado às dimensões do Aspa). A TAB. 2, a seguir, mostra os valores resultantes dessa divisão.

TABELA 2 - Frequência de ocorrência de cada padrão acentual em verbos e não verbos

Frequência de ocorrência	Posição do Acento	Verbos		Não verbos	
		N	%	N	%
Alta	Proparoxítonas	1	0,2	259	12,5
	Paroxítonas	230	50,1	1.340	64,5
	Oxítonas	228	49,7	477	23,0
<b>Total</b>		<b>459</b>	<b>100,0</b>	<b>2.076</b>	<b>100,0</b>
Baixa	Proparoxítonas	481	1,3	11.130	14,7
	Paroxítonas	27.500	71,7	50.931	67,2
	Oxítonas	10.389	27,1	13.723	18,1
<b>Total</b>		<b>38.829</b>	<b>100</b>	<b>77.860</b>	<b>100</b>

Na separação apresentada na TAB. 2, as palavras com frequência de ocorrência igual ou maior a 8007 foram consideradas como de alta frequência e as que apresentaram frequência menor que 8007 ocorrências, como de baixa frequência.

Como podemos ver, o padrão paroxítono apresenta frequência de ocorrência maior que os demais padrões. Os três padrões acentuais apresentam distribuição significativamente diferente nas categorias alta frequência ( $\chi^2[2] = 164$ ,  $p < 0,001$ ) e baixa frequência de ocorrência ( $\chi^2[2] = 5.562$ ,  $p < 0,001$ ).

Resumindo as ideias até então expostas, assume-se como hipótese de trabalho deste estudo que o acento no PB é lexicalmente especificado, o que está em consonância com a ideia de que informações redundantes estão presentes na representação mental (PIERREHUMBERT, 2001a, 2003). Defende-se que a dinâmica de atribuição acentual nesta língua pode ser bem compreendida como resultado de abstrações a partir de exemplares, as quais são responsáveis pelas tendências gerais concernentes à atribuição do acento no léxico e na gramática. Destaca-se o papel crucial dos dados experimentais a serem testados para as hipóteses aqui levantadas, que serão discutidos no cap. 6, a seguir.

Uma análise do acento no PB como a delineada neste estudo traz à tona algumas questões de extrema relevância. Dentre elas, destacam-se os seguintes tópicos: a) como uma análise estatística da fala pode oferecer generalizações relacionadas com tendências gerais da linguagem envolvendo sons e informações morfológicas; b) como se dá a relação de padrões silábicos e morfológicos com a atribuição de acento.

## 6 Análise experimental do acento

Neste capítulo será apresentada uma análise experimental do acento que busca dar suporte empírico à análise teórica apresentada nos capítulos anteriores. O capítulo se divide em três partes. A primeira apresentará considerações a respeito da metodologia experimental adotada. A segunda parte será reservada ao experimento de produção de fala; a terceira, aos três experimentos de percepção. Em cada uma das últimas partes, para cada experimento, iremos apresentar a metodologia adotada na coleta de dados (objetivos, hipóteses, estímulos, sujeitos, procedimentos), seguida dos resultados obtidos (distribuição dos dados, análise estatística, discussão dos resultados).

A análise e modelagem estatística dos dados dos quatro experimentos foi realizada por meio do R, uma linguagem e ambiente computacionais voltados para análise estatística e representação gráfica (R Development Core Team, 2011). Para maior intercâmbio de informações, os modelos testados na análise estatística serão apresentados diretamente na notação do R. O Apêndice B contém uma breve explicação de tal notação.

O protocolo de coleta de dados adotado nesta pesquisa foi aprovado pelo Comitê de Ética em Pesquisa – COEP, UFMG, no dia 12 de novembro de 2010, parecer n. ETIC 0389.0.203.000-10.

### 6.1 A metodologia experimental no estudo do acento

Um dos objetivos deste estudo é apresentar uma análise experimental do acento lexical do português brasileiro, seguindo a abordagem proposta pela Fonologia de Laboratório (PIERREHUMBERT; BECKMAN; LADD, 2000). A Fonologia de Laboratório surgiu na década de 80 como uma proposta inovadora de utilizar métodos de laboratório para analisar a estrutura sonora da linguagem humana. Contribuiu para o estabelecimento de um compromisso entre fonética-fonologia – que até então eram áreas de pesquisa independentes –, propondo uma reestruturação de suas tarefas e objetos de estudos. Essa metodologia de pesquisa tem caráter interdisciplinar, reunindo pesquisadores com trajetórias acadêmicas

variadas, comprometidos com modelos teóricos diversos e, não raro, em desacordo. Como destacam Pierrehumbert, Beckman e Ladd (2000), a Fonologia de Laboratório não se identifica com um modelo teórico específico, mas se define antes por suas opções metodológicas. Partindo da visão da linguagem humana como um fenômeno natural – e conseqüentemente da fonologia como uma ciência natural –, a Fonologia de Laboratório buscou fortalecer as bases científicas da fonologia, por meio da utilização de dados empíricos, do aprimoramento metodológico, da modelagem explícita e da acumulação de resultados.

A literatura sobre o acento no PB é composta predominantemente por estudos fonológicos tradicionais, como mostramos na seção anterior. Em outras línguas, o acento vem, cada vez mais, sendo examinado sob um ponto de vista que conjuga a análise experimental com reflexões sobre a organização gramatical (e.g. CUMMINS; PORT, 1997 e no inglês; MOOSHAMMER; FUCHS, 2002, no alemão; ORTEGA-LLEBARIA; PRIETO, 2007, no espanhol; SLUIJTER; HEUVEN, 1996, no holandês). No PB, podemos destacar os trabalhos de Barbosa (2006) e Moraes (1998) que têm conduzido estudos experimentais abrangentes sobre aspectos prosódicos do PB, abrindo caminho para pesquisas recentes que abordam o acento de forma experimental, e.g. Arantes (2010), Consoni (2006), Meireles (2007).

A análise experimental realizada neste estudo se baseou em experimentos de produção e de percepção de fala cujos resultados pudessem contribuir:

- a) com uma compreensão mais ampla dos correlatos acústicos e articulatórios do acento e sua interação com aspectos prosódicos (experimento de produção);
- b) com o debate sobre a organização e o processamento mental das informações acentuais (experimento de percepção).

## 6.2 Experimento de produção

O experimento de produção teve como objetivo avaliar as propriedades acústicas e articulatórias envolvidas na realização do acento. Pretende-se testar a hipótese de que parâmetros acústicos e articulatórios produzidos nas sílabas tônicas são diferentes das sílabas pretônicas e postônicas.

### 6.2.1 Sujeitos

Foram avaliados neste experimento cinco sujeitos, todos do sexo masculino, com nível universitário de educação formal, na faixa de 20 a 30 anos. Todos são falantes nativos do PB na variedade de Minas Gerais, não tendo se ausentado da região por mais de um ano. Nenhum deles relatou histórico de problemas de fala ou audição. Foram privilegiados sujeitos do sexo masculino, com pouca gordura e pelo facial na região do pescoço, para obter sinal eletroglotográfico de melhor qualidade, seguindo recomendação de Vieira (1997).

### 6.2.2 Procedimentos

O experimento consistiu na gravação da fala, simultaneamente à aquisição de sinal glótico e captura de movimentação da face. Para captação do sinal de fala, foi utilizado microfone de cabeça cardioide, modelo Sennheiser HSP 4. A aquisição do sinal foi realizada em um computador, por meio do *software* Audacity (AUDACITY Team, 2008), a uma taxa de amostragem de 44100Hz e a 16 bits por amostra, adequando-se à extensão em frequência dos eventos acústicos da fala.

Para captação do sinal glótico, foi utilizado um eletroglotógrafo (EGG) de dois canais, modelo EG-2 da Glottal Enterprises. O EGG consiste basicamente em dois eletrodos metálicos de contato posicionados na altura da cartilagem tireoide que medem a variação de impedância elétrica transversal resultante da abertura e fechamento das cordas vocais (VIEIRA, 1997). Quando as cordas vocais se fecham, a impedância cai, como resultado de diferenças na condutância do ar e dos tecidos da região tireoide (ROTHENBERG; MASHIE, 1988). O sinal do EGG representa a variação da área e da extensão longitudinal do contato das cordas vocais ao longo do tempo (KRISHNAMURTHY, 1983). O sinal obtido pelo EGG foi adquirido no mesmo computador e programa que o sinal de fala – cada um ocupando um canal de uma faixa estéreo.

Já a movimentação da face foi obtida pelo sistema de captura de movimentos Optotrak Certus, da Northern Digital. O Optotrak consiste em um sistema três câmeras acopladas linearmente que captam o deslocamento tridimensional de marcadores posicionados

previamente em um corpo em movimento, com resolução de até 0,1mm. O sinal do Optotrak, portanto, pode oferecer um registro preciso da movimentação de articuladores faciais potencialmente envolvidos na realização do acento, através de marcadores posicionados em diferentes regiões da face dos informantes.

A gravação foi realizada em cabine acústica do CEFALA (Centro de Estudos da Fala, Acústica, Linguagem e Música, <http://www.cefala.org>), na Escola de Engenharia da UFMG, utilizando simultaneamente três tipos de equipamento de coleta de dados: microfone, eletroglotógrafo (EGG) e sistema de captura de movimentos, como mostra o esboço do ambiente experimental mostrado na FIG. 14, a seguir. A aquisição dos sinais do EGG e da fala foi realizada no computador 1. O sinal da fala também foi capturado pelo sistema Optotrak, em módulo de aquisição próprio, simultaneamente à aquisição do movimento dos marcadores, unicamente para posterior sincronização entre sinal de EGG, sinal da fala, posição dos marcadores.

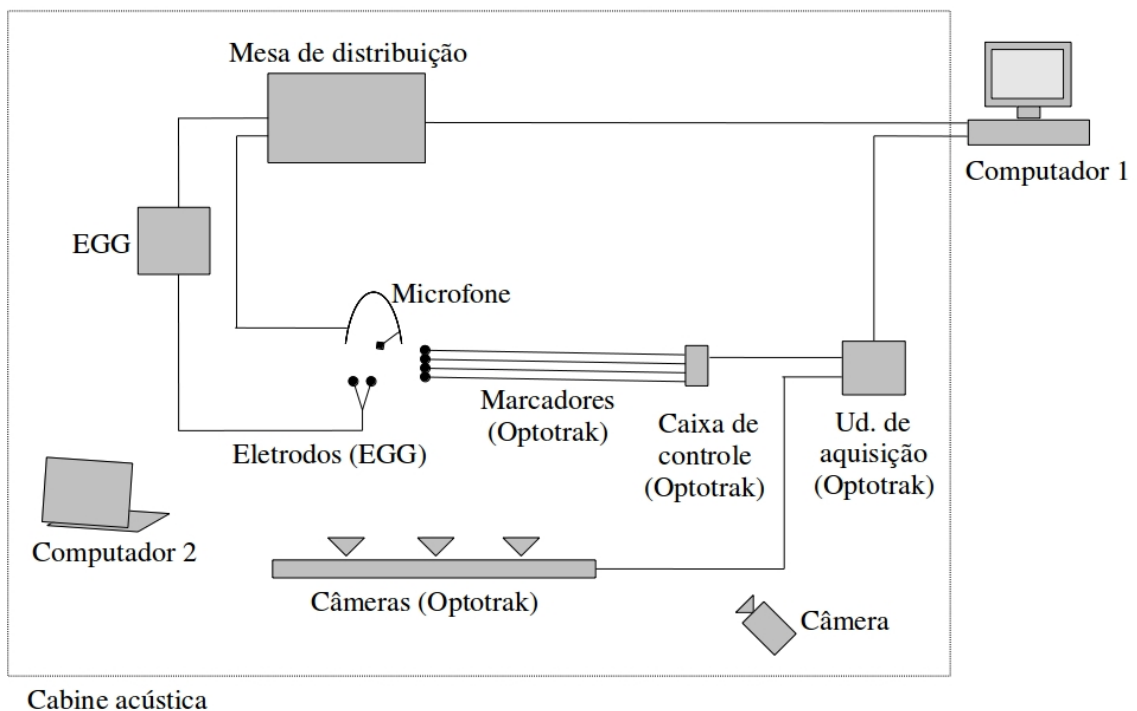


FIGURA 14 - Projeto experimental utilizado na coleta dos dados de produção

O EGG é composto por um par de eletrodos, ligados a uma unidade de processamento. O microfone capta o som da fala. O equipamento de captura de movimento Optotrak é formado por um conjunto de três câmeras alinhadas, que captam o deslocamento

tridimensional dos marcadores, estes, por sua vez, conectados a uma caixa de controle. A unidade de aquisição do Optotrak controla os marcadores por meio da caixa de controle, recebe o sinal das câmeras e gera sequências de coordenadas em três eixos ao longo do tempo. Como mostra a FIG. 14, no projeto experimental elaborado, os sinais do EGG e do microfone passam por uma mesa de distribuição, sendo, juntos, transmitidos no formato estéreo, cada um em um canal, para o computador 1, posicionado do lado de fora da cabine, para minimizar ruídos. A mesa de distribuição foi utilizada para sincronização dos sinais do EGG e da fala. O sinal do sistema Optotrak também é enviado para o computador 1, responsável pela sincronização do sinal do Optotrak com o sinal da mesa de distribuição. Os três sinais (microfone, EGG e Optotrak) foram adquiridos simultaneamente e com mesma taxa de amostragem (44100 Hz), correspondendo, portanto, aos mesmos eventos da produção oral dos sujeitos. O computador 2, portátil, foi utilizado unicamente para apresentar os estímulos a serem produzidos pelo sujeito. O experimento foi documentado em vídeo pela gravação com câmera digital.

Por restrições quanto à extensão de tempo para conclusão deste estudo, não será incluída aqui a análise dos dados do Optotrak, que deverá ser retomada no futuro. A análise dos dados de movimentação da face poderá contribuir para a compreensão dos mecanismos de coordenação de gestos e sua relação com as propriedades acústicas envolvidas na proeminência acentual.

Antes de iniciar o experimento, solicitava-se aos sujeitos que se assentassem em uma cadeira, localizada no centro da sala. Os eletrodos do EGG, presos, com uma tira, na região central do pescoço, seguindo os sinalizadores de bom posicionamento do próprio equipamento; os eletrodos eram previamente limpos e empregado gel condutivo, para melhorar a captura do sinal. Finalmente, era posicionado o microfone de cabeça, com encaixe ao redor das orelhas e a cápsula a 2 cm de distância do canto esquerdo da boca.

### 6.2.3 Estímulos

Os estímulos foram construídos de forma a possibilitar um controle rígido das variáveis independentes. Foram selecionadas quatro variáveis independentes, que passaremos a descrever.

As sílabas, ao serem comparadas com outras sílabas em uma mesma palavra, são a unidade mais relevante para o estudo do acento. Foram testadas quatro sílabas diferentes, níveis da variável que referiremos como “sílabas”: [ga], [pa], [pi] e [ku]. A escolha dos sons para formar as sílabas-teste obedeceu primeiramente a dois critérios:

- a) apenas vogais [a, i, u], visto que as médias [e, ε, o, ɔ] apresentam restrição de ocorrência em determinados ambientes (cf. SILVA, 2002);
- b) apenas consoantes oclusivas, para facilitar a segmentação – a oclusão é formada principalmente por um momento marcado em que cessa a emissão de energia acústica, resultando em uma barra vertical branca no espectrograma, o que torna a tarefa de segmentação mais acurada (LADEFOGED, 1996).

As quatro sílabas [ga], [pa], [pi] e [ku] foram especificamente selecionadas por contarem com maior número de palavras em conformidade com os critérios adotados para controle das sentenças, que serão explicitados mais adiante. Além da variável “sílabas”, outra variável independente adotada foi o tipo de relação de proeminência estabelecida entre a sílabas-teste e a sílabas acentuada da palavra. A variável “tipo” contém oito níveis:

- a) sílabas acentuada final (T1)
- b) sílabas acentuada não final seguida por uma postônica (T2)
- c) sílabas acentuada não final seguida por duas postônicas (T3)
- d) sílabas postônica final única na palavra (O4)
- e) sílabas postônica final antecedia por postônica medial (O5)
- f) sílabas postônica medial (O6)
- g) sílabas pretônica adjacente à acentuada (E7)
- h) sílabas pretônica anterior à sílabas adjacente à acentuada (E8)

A variável “tipo” foi assim construída de forma que os níveis pudessem ser facilmente agrupados quanto à relação de proeminência da sílabas teste: acentuada (T), átona postônica (O) e átona pretônica (E). Os oito níveis da variável “tipo” foram avaliados individualmente nos oito níveis listados acima, assim como agrupados pela relação de proeminência em apenas três níveis, correspondentes aos níveis marcados com T, E e O.

Conjugando as duas variáveis independentes, foram geradas 32 sílabas-teste diferentes, correspondentes a quatro níveis da variável “sílabas” (ga, pa, pi, ku) x oito níveis da variável “tipo” (T1, T2, T3, O4, O5, O6, E7, E8). A partir das 32 sílabas-teste, foram

construídas 32 palavras, por sua vez, inseridas em 32 sentenças controladas, que foram usadas como estímulos no experimento. A leitura de sentenças, ao invés de palavras isoladas, permite maior controle da estrutura prosódica dos estímulos. Ademais, procurou-se construir sentenças o mais naturais quanto possível, de forma a reduzir a artificialidade ocasionada pela experimentação em laboratório. Por essa razão, foi evitado especialmente o uso de sentenças-veículo – paradigma em que todas as palavras-teste são inseridas em uma mesma sentença –, sendo adotado em seu lugar um paradigma de sentenças controladas – ou seja, sentenças variadas, adequadas semanticamente à palavra-teste, mas controladas no nível prosódico e segmental. Na construção das sentenças, o ambiente sonoro anterior e seguinte às sílabas-teste foi controlado sistematicamente, como mostra o QUADRO 3, abaixo. Sílabas acentuadas encontram-se em negrito e sílabas-teste, sublinhadas. Asteriscos indicam logotomas de estrutura fonotática provável, criados para preencher lacunas no quadro quando não puderam ser encontradas palavras que seguissem o controle do contexto sonoro pretendido.

Tipo/Sílaba	[ga]	[pa]	[pi]	[ku]
T1	avô <b>gagá</b> tagarela	vatapá <b>garantido</b>	Pipi <b>divertida</b>	Pirarucu <b>buliçoso</b>
T2	aquela <b>gata</b> malhada	conta <b>paga</b>	Zipide*	outro <b>cu</b> bo
T3	vila <b>gátara</b> * nova	agora <b>paga</b> -se	lipídico	povo <b>cúbule</b> *
O4	<b>maga</b> taróloga	<b>mapa</b> galático	Mississippi dilúvico	<b>cuco</b> burríssimo
O5	<b>málaga</b> taça	<b>sátrapa</b> gago	<b>príncipe</b> dito	<b>búbuco</b> * bule
O6	<b>ágatas</b>	galápaga	Eurípides	<b>súcubo</b>
E7	Dinorá <b>Gatári</b> *	tafetá <b>pagável</b>	sushi <b>pedido</b>	zebu <b>cube</b> te*
E8	mesma <b>gataria</b>	mega <b>pagamento</b>	Denise <b>Pidinale</b> *	velho <b>cubulento</b> *

QUADRO 3 - Sílabas testadas e seu contexto adjacente conforme as variáveis sílabas e tipo

Como se pode verificar no QUADRO 3, os ambientes sonoros adjacentes à sílaba-teste foram controlados da seguinte forma:

- repetiu-se a mesma vogal da sílaba-teste nas sílabas anterior e seguinte;
- repetiu-se a mesma consoante inicial da sílaba-teste no início da sílaba anterior;
- utilizou-se a mesma consoante no início da sílaba seguinte, sempre uma oclusiva diferente da consoante da sílaba-teste.

Por exemplo, como se pode ver na primeira coluna do QUADRO 3, todas as palavras contendo a sílaba-teste [ga] apresentam sílaba anterior e seguinte com vogal [a], igual à sílaba-teste; sílaba anterior iniciada pela mesma consoante, [g], igual à sílaba-teste; sílaba seguinte iniciada pela mesma consoante, a oclusiva [t].

Ademais, as sentenças foram controladas quanto a:

- a) número de sílabas antes e depois da sílaba-teste, de forma a haverem sempre três sílabas antes e quatro depois;
- b) posição das outras sílabas acentuadas antes e depois da sílaba acentuada da palavra-teste, formando um pé binário antes e um terciário depois.

No Apêndice A, encontram-se listadas as 32 sentenças formuladas para o experimento. Os estímulos, em ordenação aleatória, foram apresentados por meio do *software* LibreOffice Impress, do pacote LibreOffice (2010), no computador 2. Os informantes foram solicitados a ler as 32 sentenças, apresentadas, uma após a outra, em *slides* individuais. Cada uma das sentenças foi lida em modalidade neutra e, adicionalmente, em modalidade com foco na palavra-teste. Dessa forma, foi possível comparar a realização de cada uma das sílabas-teste em estruturas com níveis de ênfase distintos (neutra x foco), a fim de se isolar características específicas da proeminência focal de características acentuais. Foi inserida no experimento, assim, uma terceira variável independente com dois níveis (neutro x foco) e que chamaremos de “foco”. A realização do foco na palavra-teste foi induzida por meio de um paradigma pergunta-resposta, em um breve treinamento que antecedeu o início do experimento. Neste treinamento, o sujeito escutava um locutor (A) dizendo uma sentença afirmativa neutra. Em seguida, escutava uma pergunta, feita por outro locutor (B), cujo questionamento dizia respeito justamente à palavra-teste. Por fim, o sujeito escutava novamente o primeiro locutor respondendo a pergunta, com ênfase na palavra-teste, elemento da sentença que era alvo do questionamento.

Exemplo do paradigma apresentado no treinamento:

(Locutor A) *Aquela casa tombada está recebendo reparos* [sentença neutra]

(Locutor B) *Aquela casa ALUGADA?*

(Locutor A) *Não, aquela casa TOMBADA está recebendo reparos.* [sentença com foco]

Este paradigma (neutra-pergunta-foco) prosseguia durante o experimento. Após o treinamento, os sujeitos eram instruídos a assumir o papel do locutor (A): primeiro deveriam pronunciar a sentença mostrada no *slide* “normalmente” (i.e. forma neutra). Depois os sujeitos ouviam o locutor (B) fazendo uma pergunta que questionava uma das palavras da sentença. Nesse momento, o controlador do experimento mudava o *slide* da apresentação, que passava a apresentar a mesma sentença com a palavra-teste em letras maiúsculas, facilitando a

identificação da palavra a receber o foco. Os sujeitos deveriam então pronunciar a sentença novamente, respondendo a pergunta (i.e., forma com foco).

Exemplo do paradigma apresentado na aquisição dos estímulos:

(*Sujeito*) *O avô gagá tagarela está dizendo bobagens* [sentença neutra]

(*Locutor B*) *O avô LELÉ tagarela?*

(*Sujeito*) *Não, o avô GAGÁ tagarela está dizendo bobagens* [sentença com foco]

As perguntas do locutor (B) ao longo do experimento, bem como as sentenças utilizadas no treinamento foram gravadas previamente por um falante natural de Belo Horizonte, do sexo masculino, com a mesma faixa etária e escolaridade dos sujeitos testados. As palavras utilizadas nas perguntas eram da mesma classe gramatical, mesmo número de sílabas e mesmo padrão acentual das palavras-teste, além de serem semanticamente plausíveis como substitutas das mesmas.

Não foram dadas instruções específicas sobre a forma com que os sujeitos deveriam responder à pergunta do locutor, pois os sujeitos demonstraram facilidade em perceber as alterações induzidas pelo foco na sentença resposta, repetindo-as naturalmente. Com adição desta variável, o número de estímulos duplicou, passando a 64 (32 sentenças x dois níveis de “foco”).

A tarefa de leitura foi repetida quatro vezes por cada informante, com cerca de três dias de diferença entre cada sessão. Foram obtidos, portanto, 1280 dados (64 sentenças x quatro sessões x cinco informantes).

#### 6.2.4 Pré-processamento dos dados

Uma vez obtidos os sinais apresentados acima, foi realizado um pré-processamento dos dados, que passaremos a descrever. O sinal de fala foi segmentado e anotado manualmente por meio da ferramenta TextGrid, do Praat, adequada à etiquetagem e anotação do sinal (BOERSMA; WEENINK, 2008). Para cada sessão de coleta, foi gerado um arquivo com extensão *wave*, estéreo, contendo o sinal do EGG no canal esquerdo e o sinal de voz, no direito. Esse arquivo foi nomeado com a inicial do informante e o número da sessão (e.g. J1

para o informante J\*\*\*, sessão 1). Foi então criado um arquivo de extensão TextGrid, pareado com o arquivo wave estéreo de cada sessão e com o mesmo nome (e.g. J1.wav; J1.TextGrid).

A FIG. 15, abaixo, mostra como exemplo um trecho da segmentação e anotação no arquivo TextGrid alinhado com o *wave* estéreo (de cima para baixo: sinal do EGG, sinal da fala, camadas de anotação/segmentação sentença, palavra, sílaba e vogal).

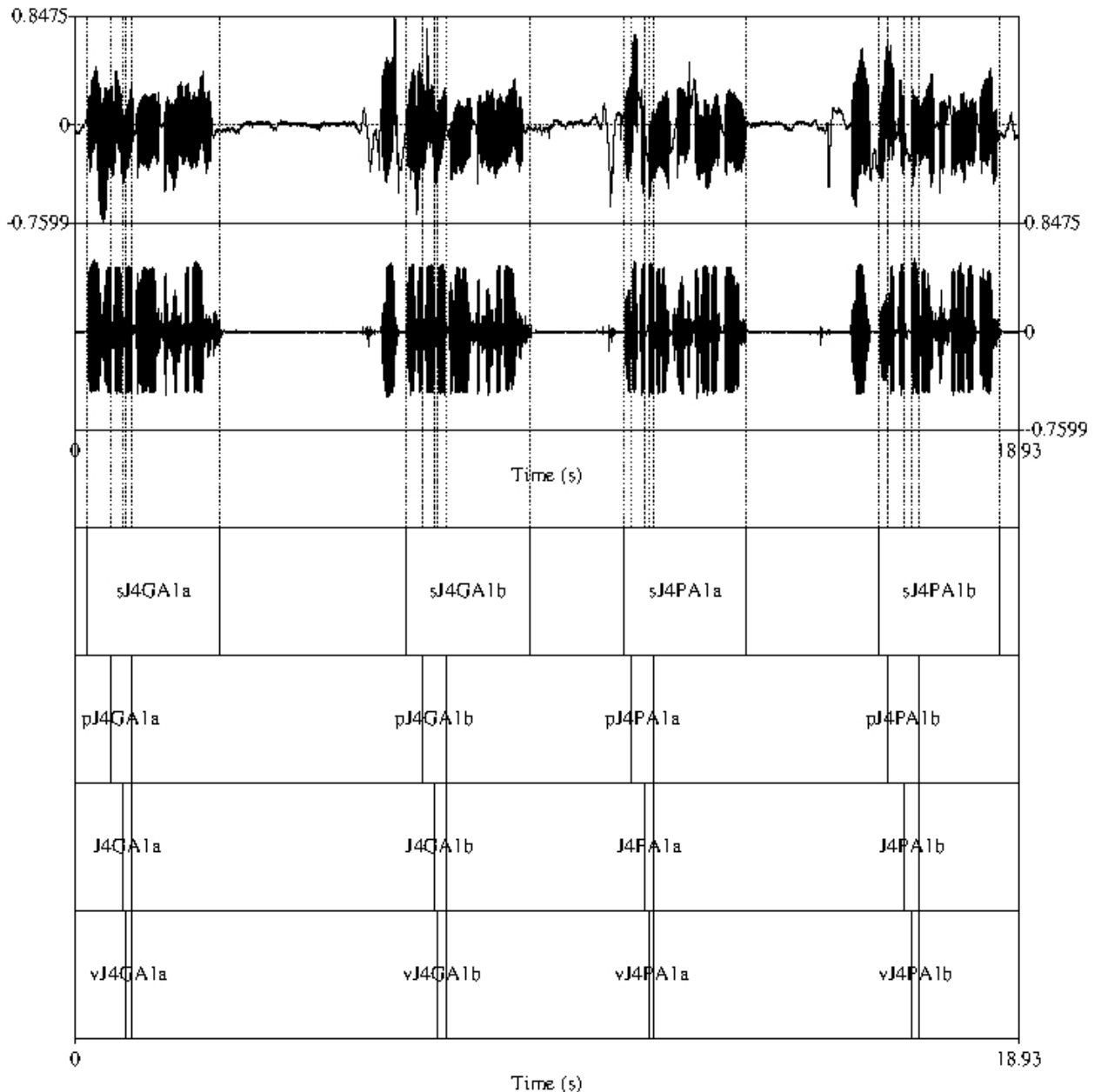


FIGURA 15 - Estrutura de camadas utilizada na segmentação do sinal obtido no experimento de produção

Como mostra a FIG. 15, a segmentação foi realizada em quatro camadas, todas de intervalo: na primeira, foram separadas as 64 sentenças testadas em cada sessão. Em seguida, em cada sentença, foi isolada a palavra-teste e, nesta, a sílaba-teste. Por fim, nas sílabas-teste, foi segmentada a vogal-teste. Dessa forma, foi obtida no TextGrid uma estrutura em quatro camadas, a serem usadas posteriormente na medição por *scripts*.

As camadas do TextGrid foram anotadas usando os códigos apresentados no QUADRO 4, que representam as variáveis independentes em estudo, acrescidas de um código para indicar o nível da segmentação em uma das quatro camadas. Um exemplo de código seria vJ4GA1a (i.e. vogal-teste, informante J\*\*\*, sessão 4, sílaba [ga], posição/distância acentuada/oxítone, neutro).

<b>Tipo</b>	<b>Símbolo</b>	<b>Significado</b>
(nível)	s	sentença
	p	palavra-teste
		sílaba-teste
	v	vogal-teste
inicial do sujeito + número da sessão de gravação	E1-E4	Sujeito E***, sessões 1 a 4
	J1-J4	Sujeito J***, sessões 1 a 4
	G1-G4	Sujeito G***, sessões 1 a 4
	R1-R4	Sujeito R***, sessões 1 a 4
	V1-V4	Sujeito V***, sessões 1 a 4
sílabas	GA	[ga]
	PA	[pa]
	PI	[pi]
	KU	[ku]
posição/distância	1	acentuada/oxítone acentuada/paroxítone
	2	acentuada/proparoxítone postônica/final
	3	
	4	adjacente postônica/medial
	5	postônica/medial não adjacente
	6	pretônica/adjacente
	7	
	8	pretônica/não adjacente
entonação	a	neutra
	b	com foco

QUADRO 4 - Códigos utilizados para anotação do sinal estéreo

A segmentação e anotação foram realizadas nos arquivos estéreo de cada sessão, em que os sinais de fala e do EGG se encontram alinhados temporalmente, exceto no caso em que era necessário segmentar uma fricativa, quando apenas o canal do sinal de fala era visualizado. A segmentação do início e fim de todos os sons foi feita com referência ao EGG, exceto no início e fim dos sons não vozeados que fossem precedidos ou seguidos por outro som não vozeado ou pausa, em que a segmentação foi feita com relação ao oscilograma ou ao espectrograma. Os seguintes ambientes sonoros foram encontrados na segmentação das sentenças, palavras-teste, sílabas-teste e vogais das sílabas-teste: início e fim de vogais orais, de consoantes oclusivas e fricativas; apenas início de consoantes africadas e nasais. Nos casos em que houve cancelamento da vogal final da sentença, considerou-se, para a segmentação, o fim da estrutura formântica das consoantes finais ou sua soltura, a depender da consoante envolvida. O QUADRO 5, a seguir, detalha os critérios de segmentação adotados para cada tipo de som e o instrumento de visualização utilizado como referência para a segmentação.

<b>Sons</b>	<b>Início</b>	<b>Fim</b>
Vogais	(EGG) primeiro cruzamento de zero positivo no primeiro ciclo glótico completo correspondente à vogal	(EGG) último cruzamento de zero negativo no último ciclo glótico completo correspondente à vogal
Sons pré- ou pós-vocálicos	(EGG) cruzamento de zero positivo após o último ciclo glótico correspondente à vogal anterior	(EGG) cruzamento de zero positivo antes do primeiro ciclo glótico correspondente à vogal seguinte
Sons precedidos ou seguidos de som não vozeado ou pausa	Oclusivas e africadas não vozeadas: (Oscilograma) início do aumento de energia correspondente à soltura Fricativas: (Espectrograma) início da região de ruído Nasais: (Espectrograma) início da barra de vozeamento	Fricativas: (Espectrograma) fim da região de ruído

QUADRO 5 - Critérios adotados na segmentação do sinal de fala

Uma vez segmentadas e anotadas as quatro camadas do TextGrid correspondente ao *wave* estéreo, realizou-se a separação dos dois canais do arquivo *wave* em dois arquivos mono, um correspondente ao sinal do EGG, outro, ao sinal de fala, por meio do *script* Monofy (ARANTES, 2010b), trabalhando ainda no Praat. Os arquivos mono foram mantidos com o mesmo nome dos estéreo, mas alocados em pastas diferentes.

Separados os dois canais, prosseguiu-se com o pré-processamento dos sinais. Quanto ao sinal do EGG, como sugerido na metodologia de Vieira (1997), foi realizada uma filtragem passa-faixa para retirada da flutuação de base no sinal, que pode ser provocada pela movimentação do pescoço, da cabeça ou dos eletrodos ou como resultado de rede elétrica. A filtragem foi realizada em MATLAB, com um filtro passa-faixas Chebyshev II 62 Hz - 7000 Hz, com *notch* em  $\approx 60$  Hz e atenuação de 45 dB na banda de rejeição. Este tipo de filtro foi selecionado por alterar em menor grau o sinal analisado, comparativamente a outros filtros, e por não introduzir *ripple* na banda de passagem, como o Chebyshev tipo I. As frequências de corte, 62 Hz a 7000Hz foram selecionadas pois se encontram acima da frequência de transmissão do sistema elétrico brasileiro e das taxas de movimentação da cabeça e laringe dos sujeitos, sendo ainda adequadas para falantes do sexo masculino. Foi utilizado um filtro passa-faixa, e não passa-altas, pois o sinal do EGG não apresenta eventos significativos acima de 7000Hz. Seguindo a metodologia de Vieira (1997), o procedimento de filtragem contemplou as seguintes etapas, com objetivo de evitar a defasagem: filtragem do sinal; inversão do sinal no eixo horizontal; filtragem do sinal invertido; re-inversão do sinal; conferência.

Após a filtragem, de volta ao Praat, o *script* Slicer (ARANTES, 2010c) foi usado para segmentar automaticamente cada vogal-teste a partir do arquivo mono filtrado com sinal de EGG. Foram gerados 1280 arquivos *wave* mono, correspondentes a cada vogal-teste, nomeados como no TextGrid (e.g. vJ4GA1a.wav). Tais arquivos foram usados nas medidas do sinal de fala, por meio de *scripts* do Praat, e do sinal do EGG, por meio do programa cedido pelo Prof. Maurílio Vieira e cujos princípios foram descritos em Vieira (1997).

### 6.2.5 Medição

Com relação ao processo de medição, iremos primeiramente explicitar as ferramentas utilizadas para realizar cada medida. Descreveremos, então, cada uma das medidas realizadas, divididas em medidas acústicas e medidas articulatórias. Por fim, retomaremos as relações entre as medidas acústicas e articulatórias apresentadas no cap. 3 e sistematizaremos possíveis relações de cada medida com a acentuação.

### *Ferramentas de medição*

O sinal de fala obtido após a fase de pré-processamento foi analisado por meio do programa Praat (BOERSMA; WEENINK, 2007). Foram utilizados *scripts* para automatizar as medidas de intensidade (ARANTES, 2008).

A partir do sinal do EGG, foram realizadas as demais medidas: duração, F0<sup>25</sup> e as medidas intrinsecamente articulatórias quociente de contato (*contact quotient*, CQ), índice de velocidade (*speed index*, SI), fase de fechamento (*closing phase*, Cp) e fase de abertura (*opening phase*, Op). Tais medidas seguem a metodologia de Vieira (1997) e foram obtidas por meio de programa cedido pelo pesquisador.

### *Medidas acústicas*

As medidas acústicas realizadas foram duração, F0 e intensidade. A duração pode ser definida como a extensão temporal de um evento – convencionalmente é dada em segundos (s), mas a unidade milissegundo (ms) é preferida por vezes, por ser mais adequada ao nível de detalhe dos eventos da fala. A intensidade, que depende da pressão sonora, corresponde à potência sonora dividida pela unidade de área e é dada em W/m<sup>2</sup>. Aqui, como na maioria dos estudos sobre fala, a medida de intensidade é dada em dB, que é uma unidade em escala logarítmica. A frequência, dada em Hz, diz respeito ao número de repetições de um padrão em um segundo. A medida de frequência que utilizaremos se refere à frequência fundamental de vibração das cordas vocais (F0).

### *Medidas articulatórias*

Foram adotadas quatro medidas articulatórias, quociente de contato (CQ), índice de velocidade (SI), fase de fechamento (Cp) e fase de abertura (Op), obtidas a partir do sinal do EGG. A FIG. 16, a seguir, reproduzida de Vieira (1987), apresenta os pontos de referência no sinal do EGG utilizados para medição do sinal glótico.

---

<sup>25</sup> Duração e F0 não são medidas intrinsecamente articulatórias, mas podem ser facilmente obtidas a partir do sinal do EGG – no caso de duração, de forma equivalente; no caso de F0, com maior precisão que se obtida a partir do sinal da fala.

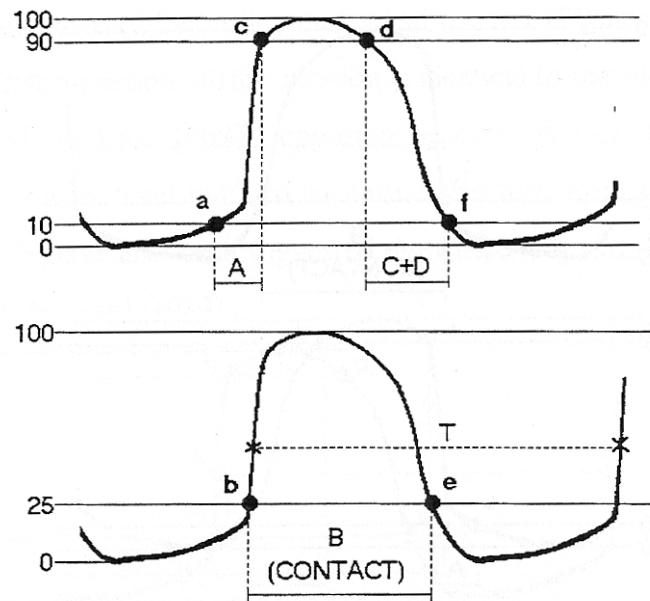


FIGURA 16 - Pontos de referência utilizados nas medidas da articulação glótica a partir do sinal do EGG

FONTE: VIEIRA, 1997, p. 105-106

Os pontos de referência indicados na FIG. 16 são definidos pelo autor como:

- a) T: período fundamental
- b) A: fase de fechamento das cordas vocais
- c) B: contato entre as cordas vocais
- d) C+D: soma das duas partes que compõem a fase de abertura das cordas vocais

A partir de um algoritmo de detecção de F0 e de tais pontos de referência, as quatro medidas específicas do sinal glótico que adotamos foram calculadas automaticamente pelo programa de Vieira (1997), da seguinte forma:

$$CQ = (B/T) \times 100\%$$

$$Cp = (A/T) \times 100\%$$

$$Op = ((C + D)/T) \times 100\%$$

$$SI = (Cp - Op)/(Cp + Op)$$

O quociente de contato (CQ), ou quociente de fechamento, corresponde à razão entre o contato das cordas vocais e o período fundamental, sendo o valor transformado em percentual. Valores de referência giram em torno de 50%.

As fases de fechamento (Cp) e de abertura (Op) foram calculadas respectivamente como a razão entre a trajetória de fechamento ou de abertura das cordas vocais e o período fundamental. Os valores são dados em porcentagens. Valores de referência giram em torno de 5 a 10% para Cp e de 33 a 39% para Op (no caso de Op, foi adotado o valor de referência obtido experimentalmente por Vieira (1997)).

O índice de velocidade (SI) é uma razão da diferença entre Cp e Op pela porção do sinal correspondente à soma das duas medidas. Varia de  $-1$  a  $1$ , sendo que o valor  $0$  corresponde a Cp e Op idênticos. O valor de referência é  $\approx -0,65$ , devido à tendência de que a fase de fechamento seja menor que a de abertura.

#### *Relações entre medidas acústicas e articulatórias*

Em geral, o aumento em F0 é acompanhado por um aumento em intensidade, mas as duas propriedades podem ser alteradas de forma independente (FANT, 1981). Considerando a alteração isolada de cada uma dessas propriedades, sabemos que F0 é diretamente proporcional à pressão subglótica, além de ser afetada pela tensão das cordas vocais. O aumento na pressão subglótica, por sua vez, pode se dar por dois mecanismos: pela ação dos pulmões e sistema respiratório ou pela resistência exercida pela glote à passagem da corrente de ar (ROTHERBERG, 1973).

Com relação à intensidade, Fant (1981) relata duas estratégias básicas: pode ser aumentada pelo aumento na pressão subglótica ou pela adução das cordas vocais causada pela compressão medial. No primeiro mecanismo, o incremento de intensidade afeta as regiões de frequência de forma aproximadamente homogênea, não sendo esperadas, portanto, alterações em ênfase espectral. Já no segundo mecanismo, a energia espectral é redistribuída e, com isso, espera-se que o aumento na intensidade seja acompanhado por um aumento em ênfase espectral e uma diminuição em Cp. Além disso, de modo geral, o aumento da intensidade levaria ao aumento de CQ, devido a um aumento do contato relativo das cordas vocais.

Espera-se que as sílabas acentuadas apresentem maior duração, maior intensidade e menor F0 (nas sentenças neutras; nas com foco, F0 seria maior nas acentuadas) que as sílabas átonas. Quanto às medidas específicas do EGG, espera-se, nas sílabas acentuadas, um maior CQ (devido a um contato vocal mais prolongado), menor SI e maior Op (devido a uma fase de abertura mais demorada) com relação às átonas.

A seguir, sistematizamos os resultados esperados das medidas acústicas e articulatórias em relação ao acento. Considerando a hipótese inicial de que parâmetros acústicos e articulatórios produzidos nas sílabas tônicas são diferentes das sílabas pretônicas e postônicas, espera-se que, em comparação com as átonas, as sílabas acentuadas tenham

a) (Medidas acústicas)

- Duração maior
- F0 menor (sentenças neutras) ou maior (nas sentenças com foco)
- Intensidade maior

b) (Medidas articulatórias)

- CQ maior
- SI menor
- Cp menor
- Op maior

Quando da análise dos dados, serão levadas em consideração diferenças inerentes às sílabas testadas – portanto, não relacionadas à acentuação e fora do campo de interesse desse estudo. Vogais altas apresentam valores mais altos de F0 que as vogais baixas, o que seria uma característica encontrada universalmente (WHALEN; LEVITT, 1995). Apenas para algumas línguas, F0 é maior nas vogais posteriores que nas anteriores. SI, ao contrário, diminui com a posterioridade da vogal (MARASEK, 1996).

## 6.2.6 Resultados

A FIG. 17, a seguir, apresenta, a título de ilustração, dois exemplos de sinais de fala e de EGG obtidos para uma sentença neutra (“Aquela gata malhada está comendo salame”) e sua versão correspondente com foco. São mostrados também a curva de F0, o espectrograma e as anotações realizadas por meio do TextGrid.

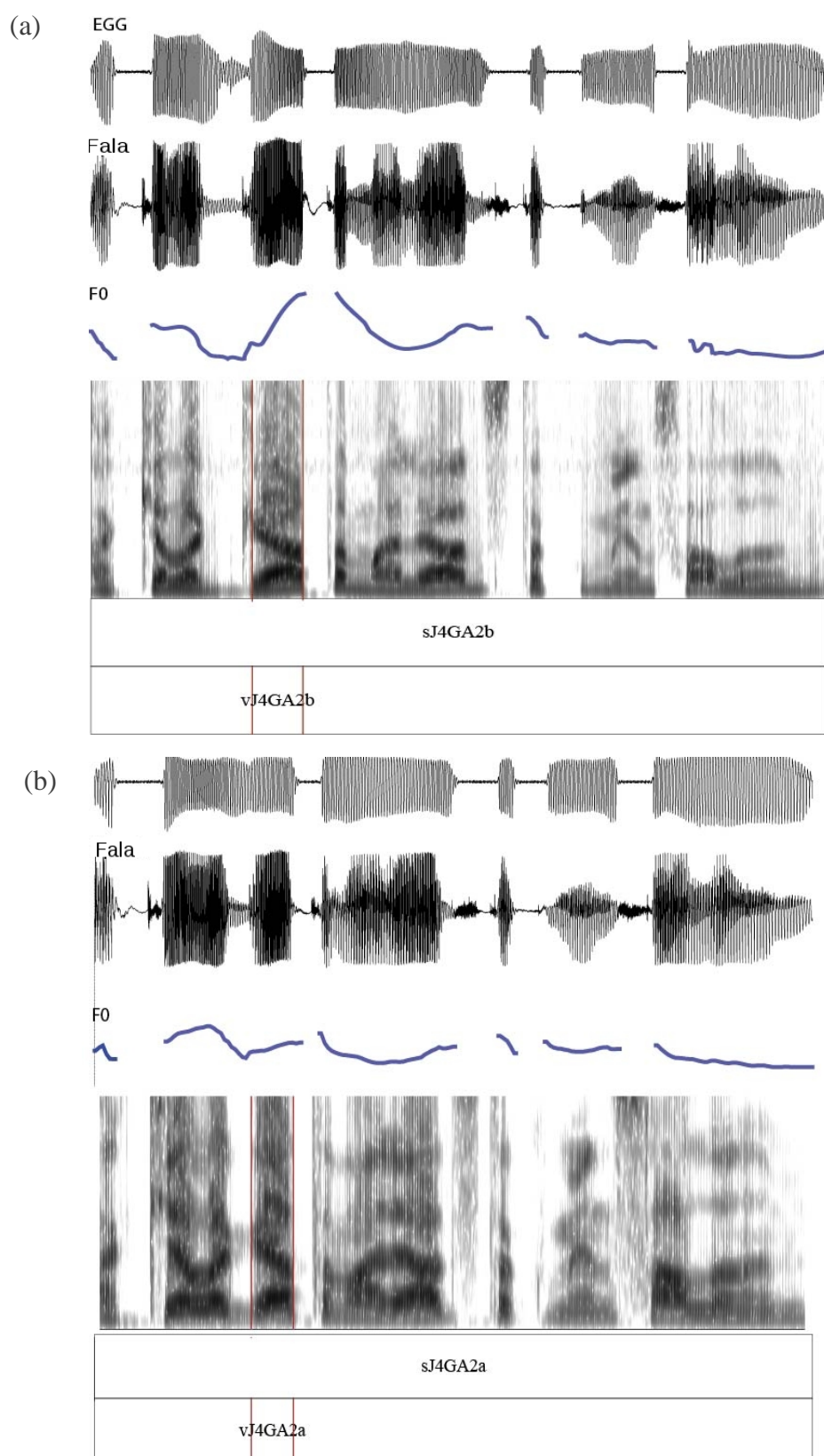


FIGURA 17 - Exemplos de dados obtidos no experimento de produção (a) modalidade neutra e (b) com foco, sentença “Aquela gata malhada está comendo salame”

Na análise dos dados, buscou-se avaliar a distribuição dos padrões acentuais por sujeito e o tipo de sílaba, assim como possíveis correlações entre as variáveis do som da fala e do sinal EGG, para cada tipo de padrão acentual.

As variáveis dependentes avaliadas neste experimento são duração (dur), intensidade (int), frequência fundamental (F0), quociente de contato (CQ), índice de velocidade (SI), fase de fechamento (CP) e fase de abertura (Op), que correspondem a parâmetros possivelmente relevantes na realização da proeminência acentual.

As variáveis independentes são:

- a) tipo: tipo de relação de proeminência entre a sílaba testada e a sílaba acentuada da palavra, com níveis T1, T2, T3, O4, O5, O6, E7, E8; foram avaliados também os agrupamentos T, O, E
- b) foco: grau de ênfase conferida à palavra da sílaba, com níveis neutro ou com foco.

Para cada uma das variáveis dependentes, foram realizados testes a fim de identificar o modelo mais adequado aos dados. Iremos descrever, como exemplo, o procedimento adotado para a variável duração. O procedimento foi realizado isoladamente para os dados referentes aos níveis neutro e com foco da variável independente foco. Note-se que, antes de iniciar a análise estatística, foi realizada, para cada item testado, uma média dos valores obtidos nas quatro sessões de gravação, para cada sujeito, para cada frase e para cada tipo de foco.

Primeiramente, os dados foram ajustados a um modelo linear<sup>26</sup> tendo como fator fixo o tipo, como mostrado em (i).

$$(i) \text{ dur} \sim \text{tipo}$$

Em seguida, foi testado um modelo linear com efeitos mistos<sup>27</sup>, em que a variável “sujeito” foi inserida como fator aleatório, como mostrado em (ii).

$$(ii) \text{ dur} \sim \text{tipo} + (1 \mid \text{sujeito})$$

Por fim, foi testado um modelo linear com efeitos mistos<sup>28</sup>, inserindo a variável “sílabas” como fator aleatório, como mostrado em (iii).

---

<sup>26</sup> R: função *lm*, pacote *stats*

<sup>27</sup> R: função *lme*, pacote *nlme*

<sup>28</sup> R: função *lmer*, pacote *lme4*

(iii)  $\text{dur} \sim \text{tipo} + (1 | \text{sujeito}) + (1 | \text{sílaba})$

Os três modelos foram comparados por um teste de razão de verossimilhança<sup>29</sup>, que respeita a distribuição qui-quadrado.

(i) vs. (ii):  $\chi^2[1] = 109,38$   $p < 0,001$

(ii) vs. (iii):  $\chi^2[1] = 260,40$   $p < 0,001$

A diferença entre os modelos é estatisticamente significativa, o que justifica a utilização do modelo com os dois fatores aleatórios, que, apesar de apresentar mais coeficientes, se ajusta melhor aos dados que os demais.

O mesmo procedimento foi seguido para cada uma das variáveis dependentes: além de duração, F0, intensidade, CQ, SI, Cp, Op. Para todas elas, o modelo misto com dois fatores aleatórios se mostrou mais adequado e foi, portanto, o adotado na análise.

Identificado o modelo ao qual ajustar os dados, para cada variável dependente, realizamos testes simultâneos de comparações múltiplas de médias<sup>30</sup>, para determinar a significância da interação entre os níveis das variáveis independentes inseridas como fatores fixos no modelo.

Na avaliação dos resultados, que passaremos a realizar, apresentaremos, para cada variável dependente, um gráfico com a distribuição dos dados, seguido de observações gerais e do teste de comparações múltiplas, que determina a significância de algumas das tendências observadas na distribuição dos dados.

Nas FIG. 18 a 24, a seguir, é mostrada, em *boxplots*, a distribuição dos dados obtidos no experimento para cada variável dependente, disposta no eixo vertical dos gráficos. No eixo horizontal, aparecem as variáveis tipo de proeminência, com seus oito níveis (numerados de 1 a 8), e acento, em dois níveis (a = neutro e b = foco). As cores das caixas dos gráficos correspondem ao agrupamento da variável tipo de proeminência em três níveis: acentuada (T) = rosa, postônica (O) = azul e pretônica (E) = amarelo.

Segundo Crawley (2007), os *boxplots* mostram a distribuição dos dados quanto a sua localização e dispersão, assim como a assimetria da distribuição. Em cada retângulo do gráfico, os limites inferior e superior correspondem respectivamente ao primeiro e ao terceiro quartis, representando, portanto, como estão distribuídos os 50% centrais dos dados. A linha

<sup>29</sup> R: função *anova*, pacote *stats*

<sup>30</sup> R: função *glht*, pacote *multcomp*

mais escura no interior da caixa indica a mediana ou o centro dos dados. O deslocamento do centro dos dados (mediana) com relação ao centro do retângulo corresponde à assimetria dos dados. As linhas tracejadas mostram o que for menor entre o valor máximo dos dados e 1,5 vezes a distância entre o primeiro e terceiro quartis (aproximadamente dois desvios-padrão), representando a variabilidade dos dados. Os eventuais pontos que se estendem além da linha tracejada são *outliers*. Adicionalmente, incluímos nos *boxplots* a média, representada por um círculo sólido cinza, e o erro padrão, representado por um traço sólido cinza, ambos sobrepostos aos retângulos de cada grupo.

Depois de cada *boxplot*, serão apresentadas tabelas com resultados de teste estatístico, com a finalidade de determinar se as distribuições dos dados observadas no gráfico são significativamente diferentes umas das outras. A estatística utilizada em tais comparações é o *z-score*, que é um teste de hipótese baseado no desvio padrão. Testa se as médias de duas distribuições são iguais (hipótese nula). De forma simplificada, o valor de *p* corresponde à probabilidade de que a hipótese nula seja verdadeira, considerando o *z-score* encontrado. Adotaremos significância a  $p \leq \alpha = 0,05$  (que equivale a um nível de confiança de 95%).

A seguir, a FIG. 18 mostra a distribuição dos dados para a variável duração.

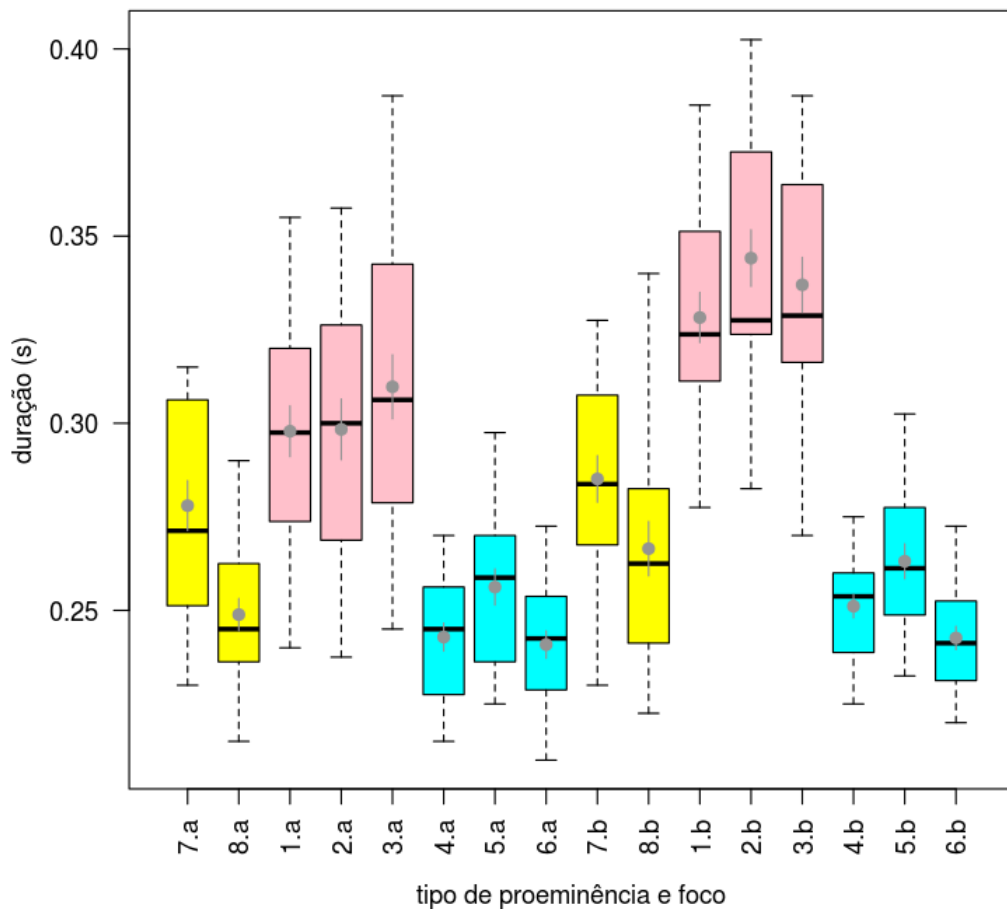


FIGURA 18 - Distribuição dos dados de duração (s) de acordo com tipo de proeminência, localização da proeminência e foco

<i>Tipo de proeminência</i>	<i>Localização da proeminência</i>
<span style="color: #f08080;">■</span> acentuada	1 acentuada final
<span style="color: #00bfff;">■</span> postônica	2 acentuada seguida por uma postônica
<span style="color: #ffff00;">■</span> pretônica	3 acentuada seguida por duas postônicas
	4 postônica final única
	5 postônica final antecedida por postônica medial
<i>Foco</i>	6 postônica medial
a neutra	7 pretônica adjacente à acentuada
b com foco	8 pretônica anterior à sílaba adjacente à acentuada

Na FIG. 18, podemos identificar as seguintes características de distribuição da propriedade duração: as sílabas acentuadas (rosa) apresentam maior variabilidade que as pretônicas (amarelo) e estas, por sua vez, que as postônicas (azul). Há grande assimetria nas acentuadas com foco (1b, 2b e 3b). São muito próximos os centros das acentuadas neutras entre si; das acentuadas com foco entre si; das postônicas final única e medial entre si, o que

seria uma pista de que as amostras não são significativamente diferentes umas das outras.

Podem ser observadas tendências claras com relação aos valores de duração: as sílabas acentuadas apresentam maior duração que as átonas, tanto nas sentenças neutras, quanto nas com foco. Entre as átonas, as sílabas pretônicas se mostraram mais longas que as postônicas.

Essas observações gerais foram avaliadas por testes simultâneos de comparações múltiplas de médias, que são mostrados na TAB. 3, a seguir.

TABELA 3 - Resultados dos testes de comparações múltiplas para a variável duração

Duração							
neutro				foco			
Níveis	valor z	p (> z )		Níveis	valor z	p (> z )	
T ~ O	18,915	<1e-04	***	T ~ O	28,901	<1e-04	***
T ~ E	11,339	<1e-04	***	T ~ E	18,154	<1e-04	***
O ~ E	-9,657	<1e-04	***	O ~ E	-14,079	<1e-04	***
T1 ~ T2	-0,099	1,000		T1 ~ T2	-3,147	0,0139	*
T1 ~ T3	-2,344	0,140		T1 ~ T3	-1,735	0,4641	
T2 ~ T3	-2,245	0,176		T2 ~ T3	1,413	0,6952	
O4 ~ O5	-2,640	0,065		O4 ~ O5	-2,379	0,1286	
O4 ~ O6	0,395	0,999		O4 ~ O6	1,685	0,4993	
O5 ~ O6	3,034	0,020	*	O5 ~ O6	4,064	0,0004	***
E7 ~ E8	-2,640	0,065		E7 ~ E8	-2,379	0,1286	

Códigos de significância: \*\*\* 0,001 \*\* 0,01 \* 0,05

A TAB. 3 mostra que, nas sentenças neutras – considerando a significância a  $p \leq 0,05$  –, a variável duração se mostrou significativa para diferenciar acentuadas (T) de postônicas (O), acentuadas (T) de pretônicas (E), postônicas (O) de pretônicas (E), assim como as sílabas postônicas finais (O5) e mediais (O6) em uma proparoxítonea.

Nas sentenças com foco, foram observados resultados semelhantes, exceto por, nesse caso, a duração ter adicionalmente se mostrado significativa para diferenciar sílabas acentuadas finais (T1) de sílabas acentuadas seguidas por uma postônica (T2).

A seguir, a FIG. 19 mostra a distribuição dos dados para a variável F0.

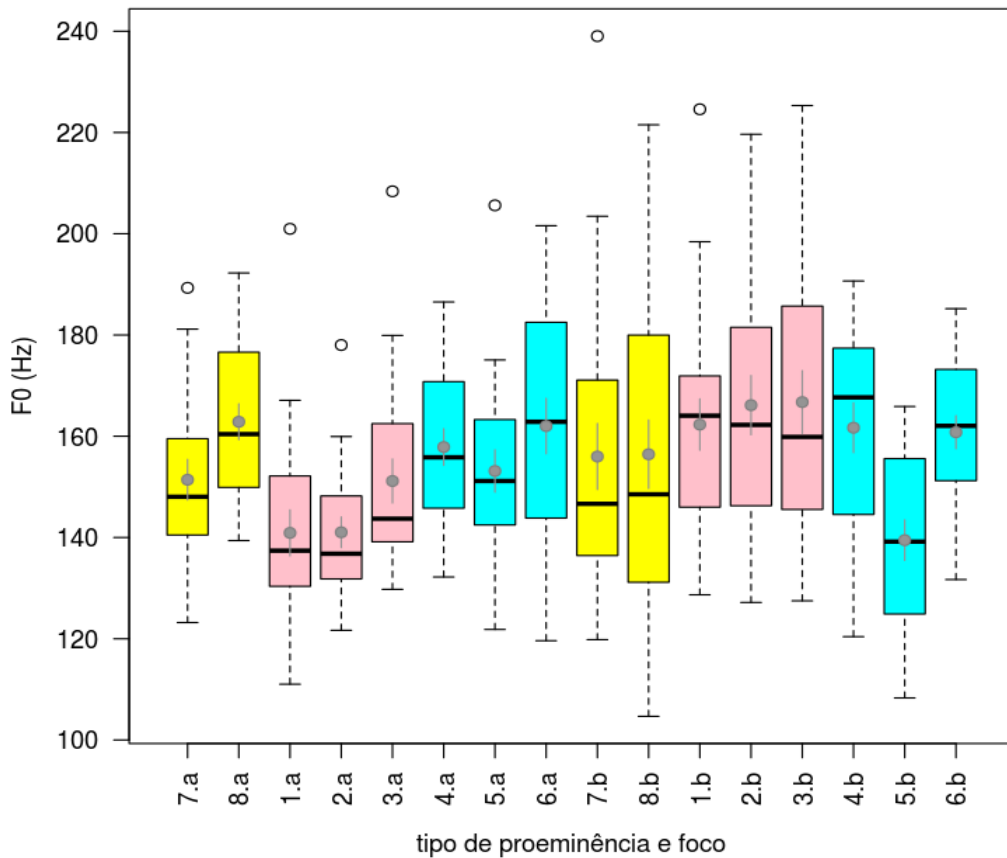


FIGURA 19 - Distribuição dos dados de F0 (Hz) de acordo com tipo de proeminência, localização da proeminência e foco

<i>Tipo de proeminência</i>	<i>Localização da proeminência</i>
<span style="display: inline-block; width: 15px; height: 15px; background-color: #FF9999; border: 1px solid black;"></span> acentuada	1 acentuada final
<span style="display: inline-block; width: 15px; height: 15px; background-color: #00FFFF; border: 1px solid black;"></span> postônica	2 acentuada seguida por uma postônica
<span style="display: inline-block; width: 15px; height: 15px; background-color: #FFFF00; border: 1px solid black;"></span> pretônica	3 acentuada seguida por duas postônicas
	4 postônica final única
	5 postônica final antecedida por postônica medial
<i>Foco</i>	6 postônica medial
a neutra	7 pretônica adjacente à acentuada
b com foco	8 pretônica anterior à sílaba adjacente à acentuada

Na FIG. 19, podemos identificar as seguintes características de distribuição da propriedade F0: há grande variabilidade em algumas categorias, mas aparentemente assistemática. Há certa assimetria nas acentuadas neutras (1a, 2a e 3a) e, nas sentenças com foco, na acentuada final (1b). São muito próximos os centros das acentuadas neutras 1a e 2a entre si; das acentuadas com foco entre si; das postônicas final única (4a) e final antecedida

por medial (5a) entre si, o que seria uma pista de que tais amostras não são significativamente diferentes entre si.

Podem ser observadas algumas tendências com relação aos valores de F0: as sílabas acentuadas apresentam menor F0 que as átonas nas sentenças neutras, o que se inverte nas sentenças com foco, em que as sílabas acentuadas apresentam maiores valores de F0 que as átonas. Essas observações gerais foram avaliadas por testes simultâneos de comparações múltiplas de médias, que são mostrados na TAB. 4, a seguir.

TABELA 4 - Resultados dos testes de comparações múltiplas para a variável F0

F0							
neutro				foco			
Níveis	valor z	p (> z )		Níveis	valor z	p (> z )	
T ~ O	-6,567	<1e-04	***	T ~ O	3,221	0,011	*
T ~ E	-4,449	<1e-04	***	T ~ E	2,427	0,114	
O ~ E	2,934	0,028	*	O ~ E	-1,240	0,806	
T1 ~ T2	-0,033	1,000		T1 ~ T2	-0,642	0,991	
T1 ~ T3	-2,919	0,029	*	T1 ~ T3	-0,743	0,980	
T2 ~ T3	-2,886	0,032	*	T2 ~ T3	-0,101	1,000	
O4 ~ O5	1,348	0,739		O4 ~ O5	3,726	0,002	**
O4 ~ O6	-1,183	0,838		O4 ~ O6	0,149	1,000	
O5 ~ O6	-2,530	0,088		O5 ~ O6	-3,577	0,003	**
E7 ~ E8	1,348	0,739		E7 ~ E8	3,726	0,002	**

Códigos de significância: \*\*\* 0,001 \*\* 0,01 \* 0,05

A TAB. 4 mostra que, nas sentenças neutras – considerando a significância a  $p \leq 0,05$  –, a variável F0 se mostrou significativa para diferenciar acentuadas (T) de postônicas (O), acentuadas (T) de pretônicas (E), postônicas (O) de pretônicas (E), assim como para diferenciar acentuadas finais (T1) de acentuadas seguidas por duas postônicas (T3) e acentuadas seguidas por uma postônica (T2) de acentuadas seguidas por duas postônicas (T3).

Nas sentenças com foco, foram observados resultados muito diferentes dos observados nas sentenças neutras: F0 mostrou-se relevante para distinguir apenas as sílabas postônicas únicas (O4) de postônicas finais antecidas de postônica medial (O5), estas de postônicas mediais (O6) e pretônicas adjacentes (E7) de pretônicas anteriores à adjacente (E8).

A seguir, a FIG. 20 mostra a distribuição dos dados para a variável intensidade.

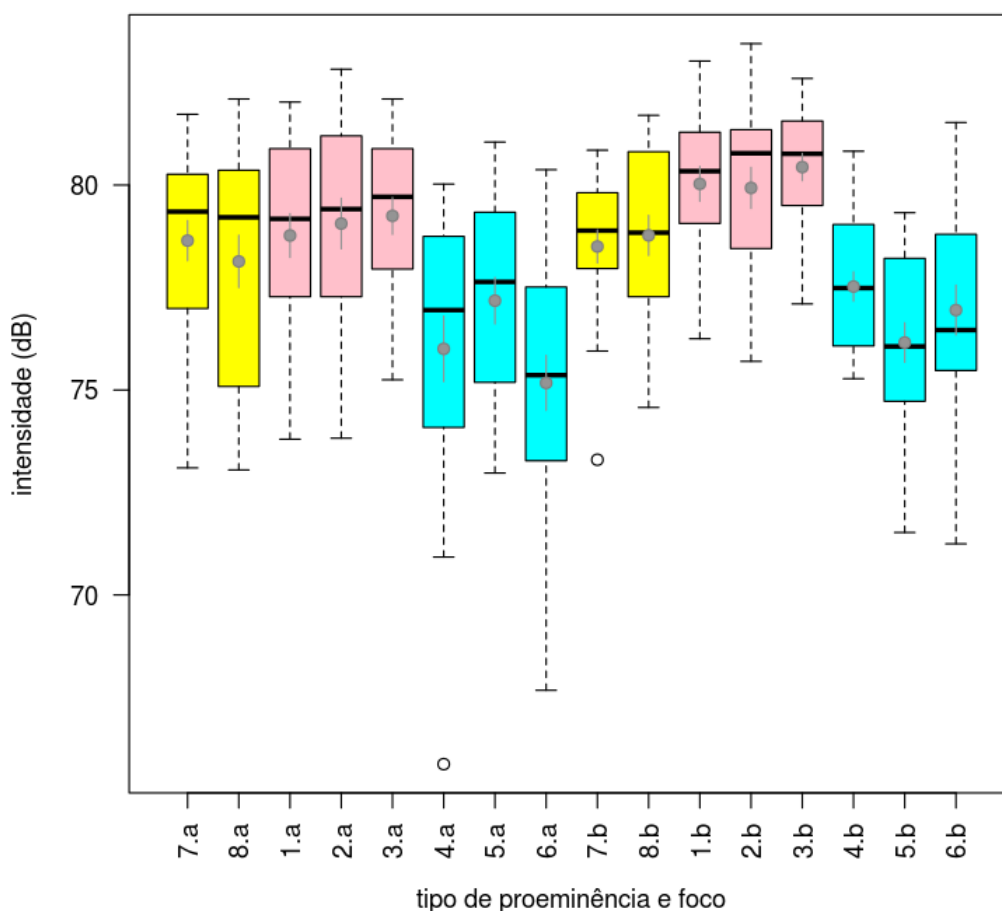


FIGURA 20 - Distribuição dos dados de intensidade (dB) de acordo com tipo de proeminência, localização da proeminência e foco

<i>Tipo de proeminência</i>	<i>Localização da proeminência</i>
■ acentuada	1 acentuada final
■ postônica	2 acentuada seguida por uma postônica
■ pretônica	3 acentuada seguida por duas postônicas
	4 postônica final única
	5 postônica final antecedida por postônica medial
<i>Foco</i>	6 postônica medial
a neutra	7 pretônica adjacente à acentuada
b com foco	8 pretônica anterior à sílaba adjacente à acentuada

Na FIG. 20, podemos identificar as seguintes características de distribuição da propriedade intensidade: destaca-se das demais a variabilidade da postônica medial neutra (6a). Há assimetria nas pretônicas neutras (7a, 8a) e, nas sentenças com foco, na acentuada seguida por uma postônica (2b). São muito próximos os centros das acentuadas neutras entre si; das pretônicas neutras entre si; das acentuadas com foco entre si; das pretônicas com foco

entre si, o que seria uma pista de que tais amostras não são significativamente diferentes entre si.

Podem ser observadas as seguintes tendências com relação aos valores de intensidade: tanto as sílabas acentuadas quanto as pretônicas apresentam maior intensidade que as postônicas, o que acontece nas sentenças neutras e com foco.

Essas observações gerais foram avaliadas por testes simultâneos de comparações múltiplas de médias, que são mostrados na TAB. 5, a seguir.

TABELA 5 - Resultados dos testes de comparações múltiplas para a variável intensidade

Intensidade							
neutro				foco			
Níveis	valor z	p (> z )		Níveis	valor z	p (> z )	
T ~ O	8,295	<1e-04	***	T ~ O	10,861	< 1e-04	***
T ~ E	1,779	0,433		T ~ E	4,218	0,0002	***
O ~ E	-6,842	<1e-04	***	O ~ E	-7,417	< 1e-04	***
T1 ~ T2	-0,480	0,998		T1 ~ T2	0,190	1,0000	
T1 ~ T3	-0,789	0,974		T1 ~ T3	-0,787	0,9738	
T2 ~ T3	-0,309	1,000		T2 ~ T3	-0,978	0,9270	
O4 ~ O5	-1,936	0,332		O4 ~ O5	2,632	0,0667	
O4 ~ O6	1,369	0,725		O4 ~ O6	1,105	0,8763	
O5 ~ O6	3,305	0,008	**	O5 ~ O6	-1,527	0,6141	
E7 ~ E8	-1,936	0,332		E7 ~ E8	2,632	0,0666	

Códigos de significância: \*\*\* 0,001 \*\* 0,01 \* 0,05

A TAB. 5 mostra que, nas sentenças neutras – considerando a significância a  $p \leq 0,05$  –, a variável intensidade se mostrou significativa para diferenciar acentuadas (T) de postônicas (O) e postônicas (O) de pretônicas (E), assim como para diferenciar postônicas finais antecedidas de postônica medial (O5) de postônicas mediais (O6).

Nas sentenças com foco, foram observadas algumas diferenças: intensidade mostrou-se significativa para diferenciar acentuadas (T) de postônicas (O), acentuadas (T) de pretônicas (E), postônicas (O) de pretônicas (E).

A seguir, a FIG. 21 mostra a distribuição dos dados para a variável CQ.

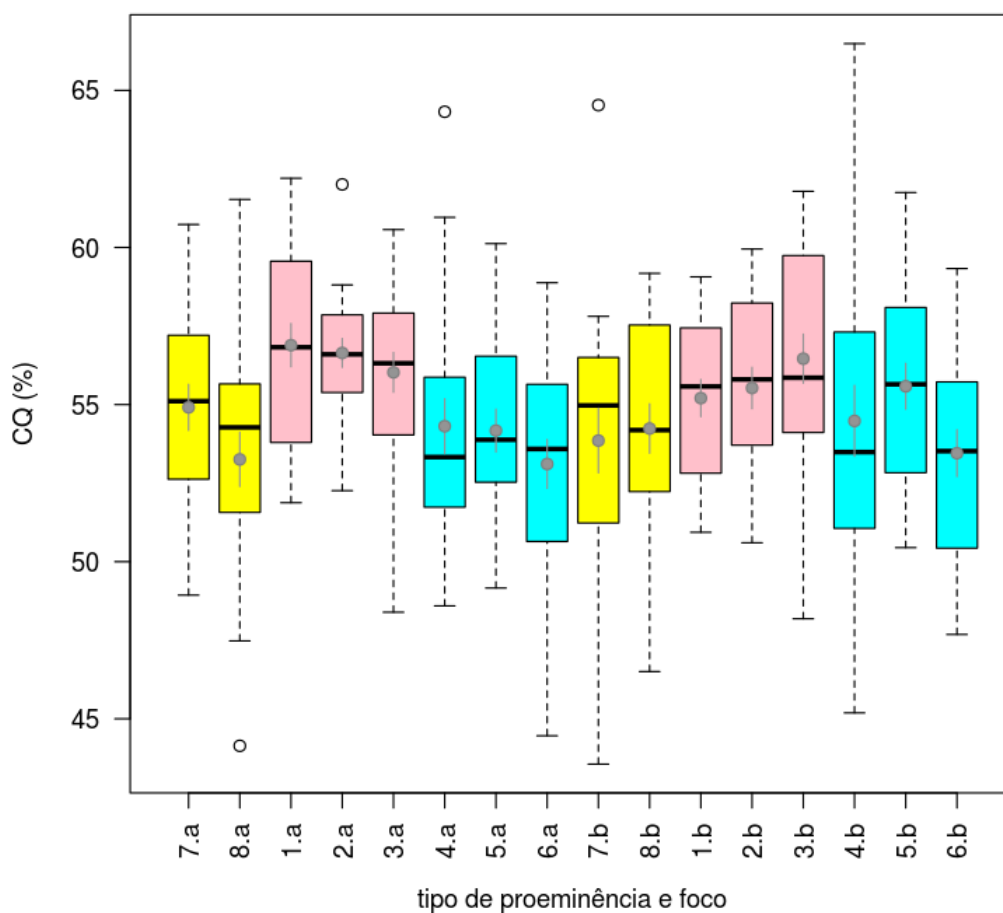


FIGURA 21 - Distribuição dos dados de quociente de contato CQ (%) de acordo com tipo de proeminência, localização da proeminência e foco

<i>Tipo de proeminência</i>	<i>Localização da proeminência</i>
acentuada	1 acentuada final
postônica	2 acentuada seguida por uma postônica
pretônica	3 acentuada seguida por duas postônicas
	4 postônica final única
	5 postônica final antecedida por postônica medial
	6 postônica medial
	7 pretônica adjacente à acentuada
	8 pretônica anterior à sílaba adjacente à acentuada

<i>Foco</i>
a neutra
b com foco

Na FIG. 21, podemos identificar as seguintes características de distribuição da propriedade quociente de contato (CQ): há grande variabilidade em todas as categorias, exceto, nas neutras, a acentuada seguida de uma postônica (2a) e, nas com foco, a acentuada final (1b) e também a acentuada seguida de uma postônica (2b). Há certa assimetria, nas

sentenças neutras, na postônica final (4a), na postônica antecedida de postônica medial (5a) e na pretônica não adjacente (8a). Nas sentenças com foco, há assimetria na acentuada seguida de duas postônicas (3b), bem como nas pretônicas. São muito próximos os centros das acentuadas neutras entre si; das postônicas neutras entre si; das pretônicas neutras entre si; das acentuadas com foco entre si, o que seria uma pista de que tais amostras não são significativamente diferentes entre si.

Essas observações gerais foram avaliadas por testes simultâneos de comparações múltiplas de médias, que são mostrados na TAB. 6, a seguir.

TABELA 6 - Resultados dos testes de comparações múltiplas para a variável CQ

CQ						
neutro			foco			
Níveis	valor z	p (> z )	Níveis	valor z	p (> z )	
T ~ O	4,681	<1e-04	***	T ~ O	1,932	0,334
T ~ E	3,241	0,010	*	T ~ E	2,505	0,094
O ~ E	-2,035	0,275		O ~ E	0,113	1,000
T1 ~ T2	0,253	1,000		T1 ~ T2	-0,293	1,000
T1 ~ T3	0,881	0,955		T1 ~ T3	-1,141	0,859
T2 ~ T3	0,628	0,992		T2 ~ T3	-0,848	0,962
O4 ~ O5	0,143	1,000		O4 ~ O5	-1,002	0,919
O4 ~ O6	1,225	0,815		O4 ~ O6	0,934	0,941
O5 ~ O6	1,081	0,887		O5 ~ O6	1,937	0,331
E7 ~ E8	0,143	1,000		E7 ~ E8	-1,002	0,918

Códigos de significância: \*\*\* 0,001 \*\* 0,01 \* 0,05

Como podemos ver na TAB. 6, nas sentenças neutras – considerando a significância a  $p \leq 0,05$  –, a variável CQ se mostrou significativa para diferenciar acentuadas (T) de postônicas (O) e postônicas (O) de pretônicas (E), assim como para diferenciar postônicas finais antecedidas de postônica medial (O5) de postônicas mediais (O6).

Nas sentenças com foco, foram observadas algumas diferenças: CQ mostrou-se significativo para diferenciar acentuadas (T) de postônicas (O), acentuadas (T) de pretônicas (E), postônicas (O) de pretônicas (E).

A seguir, a FIG. 22 mostra a distribuição dos dados para a variável SI.

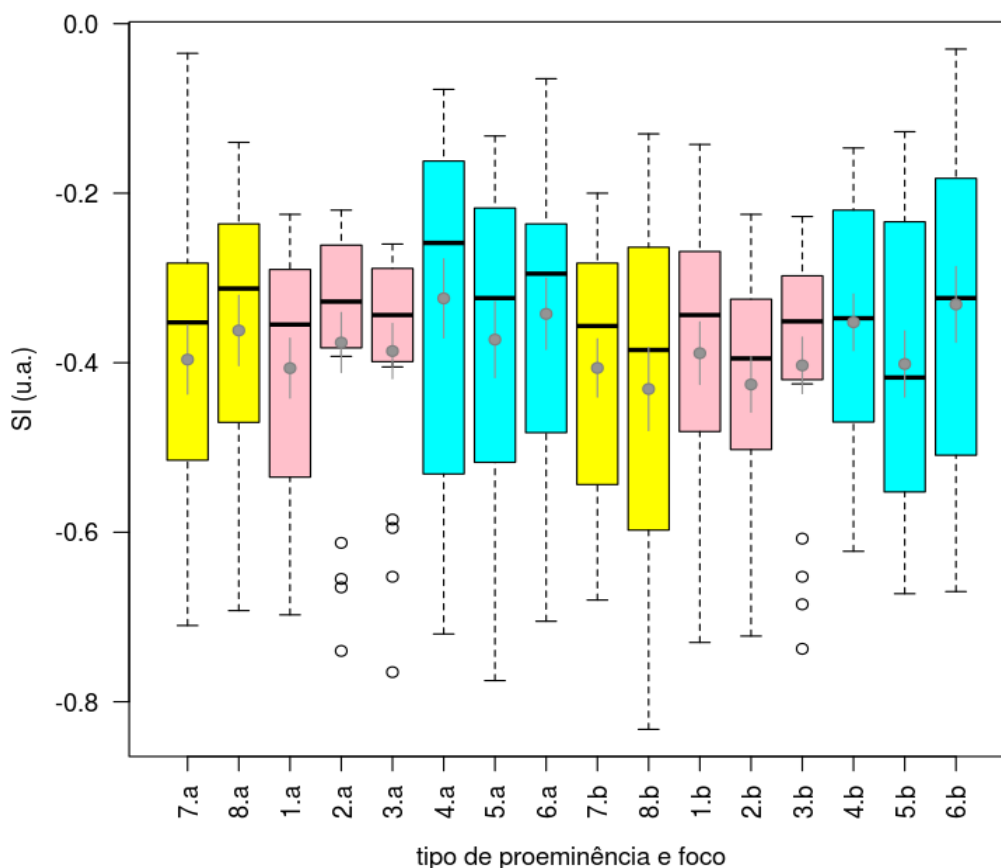


FIGURA 22 - Distribuição dos dados de índice de velocidade SI (u.a.) de acordo com tipo de proeminência, localização da proeminência e foco

<i>Tipo de proeminência</i>	<i>Localização da proeminência</i>
acentuada	1 acentuada final
postônica	2 acentuada seguida por uma postônica
pretônica	3 acentuada seguida por duas postônicas
	4 postônica final única
	5 postônica final antecedida por postônica medial
	6 postônica medial
	7 pretônica adjacente à acentuada
	8 pretônica anterior à sílaba adjacente à acentuada

Na FIG. 22, podemos identificar as seguintes características de distribuição da propriedade índice de velocidade (SI): há grande variabilidade em todas as categorias, exceto nas acentuadas neutras seguidas de uma postônica (2a) e de duas postônicas (3a), assim como na acentuada com foco seguida de duas postônicas (3b). Nesses três casos, note-se, contudo, a presença de *outliers* no que corresponde à cauda inferior dos dados. A assimetria pode ser

descrita de forma semelhante, ocorrendo de forma generalizada nas categorias, exceto nos três casos de pouca variabilidade, mas com *outliers*. Quanto aos centros, aproximam-se as acentuadas neutras seguidas de uma (2a) e duas postônicas (3a); as postônicas neutras final (4a) e medial (6a); nas sentenças com foco, a acentuada final (1b) e seguida de duas postônicas. A proximidade do centro dessas amostras seria uma pista de que as mesmas não são significativamente diferentes entre si.

Essas observações gerais foram avaliadas por testes simultâneos de comparações múltiplas de médias, que são mostrados na TAB. 7, a seguir.

TABELA 7 - Resultados dos testes de comparações múltiplas para a variável SI

SI					
neutro			foco		
Níveis	valor z	p (> z )	Níveis	valor z	p (> z )
T ~ O	-2,176	0,205	T ~ O	-2,497	0,096
T ~ E	-0,118	1,000	T ~ E	0,269	1,000
O ~ E	2,070	0,256	O ~ E	2,706	0,054
T1 ~ T2	-0,998	0,920	T1 ~ T2	1,203	0,827
T1 ~ T3	-0,665	0,989	T1 ~ T3	0,469	0,998
T2 ~ T3	0,333	1,000	T2 ~ T3	-0,734	0,982
O4 ~ O5	1,606	0,557	O4 ~ O5	1,605	0,558
O4 ~ O6	1,089	0,884	O4 ~ O6	-0,688	0,987
O5 ~ O6	-0,495	0,998	O5 ~ O6	-2,292	0,158
E7 ~ E8	1,606	0,557	E7 ~ E8	1,605	0,558

Códigos de significância: \*\*\* 0,001    \*\* 0,01    \* 0,05

Como podemos ver na TAB. 7, de modo geral – considerando a significância a  $p \leq 0,05$  –, a variável SI não se mostrou significativa em nenhuma das comparações. Esse resultado pode estar relacionado à grande variabilidade observada em todas as categorias.

A seguir, a FIG. 23 mostra a distribuição dos dados para a variável Cp.

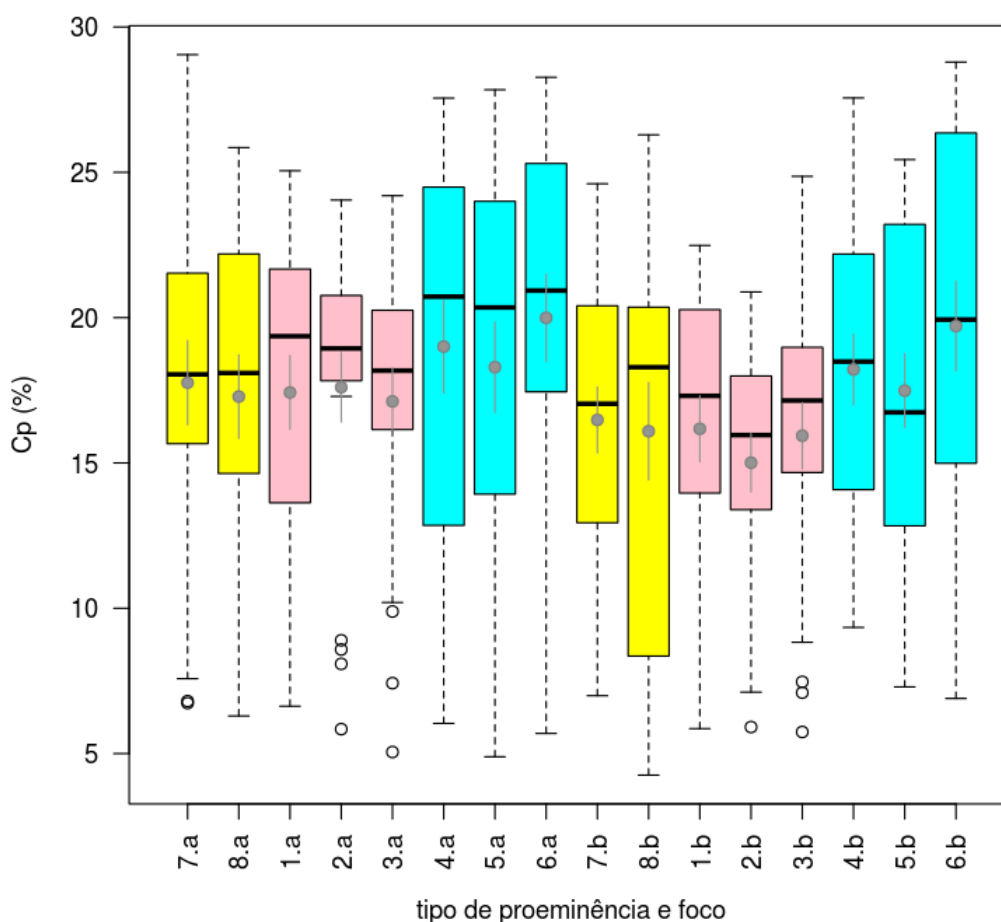


FIGURA 23 - Distribuição dos dados de fase de fechamento Cp (%) de acordo com tipo de proeminência, localização da proeminência e foco

<i>Tipo de proeminência</i>	<i>Localização da proeminência</i>
acentuada	1 acentuada final
postônica	2 acentuada seguida por uma postônica
pretônica	3 acentuada seguida por duas postônicas
	4 postônica final única
	5 postônica final antecedida por postônica medial
	6 postônica medial
	7 pretônica adjacente à acentuada
	8 pretônica anterior à sílaba adjacente à acentuada

<i>Foco</i>
a neutra
b com foco

Na FIG. 23, podemos identificar as seguintes características de distribuição da propriedade fase de fechamento (Cp): há grande variabilidade em todas as categorias, exceto na acentuada neutra seguida uma postônica (2a), em que há, contudo, muitos *outliers* no que corresponde à cauda esquerda dos dados. Há assimetria na maioria das categorias, exceto na acentuada neutra seguida de duas postônicas (3a) e, nas sentenças com foco, nas acentuadas,

na postônica final (4b) e na pretônica não adjacente (7b). Quanto aos centros, aproximam-se as postônicas neutras; as pretônicas neutras; nas sentenças com foco, as acentuadas final (1b) e seguida de duas postônicas (3b).

Essas observações gerais foram avaliadas por testes simultâneos de comparações múltiplas de médias, que são mostrados na TAB. 8, a seguir.

TABELA 8 - Resultados dos testes de comparações múltiplas para a variável Cp

Cp						
neutro			foco			
Níveis	valor z	p (> z )	Níveis	valor z	p (> z )	
T ~ O	-2,924	0,028	T ~ O	-4,916	< 1e-04	***
T ~ E	-0,311	1,000	T ~ E	-1,120	0,869	
O ~ E	2,657	0,062	O ~ E	3,978	0,001	***
T1 ~ T2	-0,199	1,000	T1 ~ T2	1,199	0,829	
T1 ~ T3	0,338	1,000	T1 ~ T3	0,243	1,000	
T2 ~ T3	0,537	0,997	T2 ~ T3	-0,956	0,934	
O4 ~ O5	0,782	0,975	O4 ~ O5	0,750	0,979	
O4 ~ O6	-0,491	0,998	O4 ~ O6	-1,533	0,610	
O5 ~ O6	-1,263	0,792	O5 ~ O6	-2,283	0,162	
E7 ~ E8	0,782	0,975	E7 ~ E8	0,750	0,979	

Códigos de significância: \*\*\* 0,001 \*\* 0,01 \* 0,05

Como podemos ver na TAB. 8, nas sentenças neutras – considerando a significância a  $p \leq 0,05$  –, a variável Cp se mostrou significativa apenas para diferenciar acentuadas (T) de postônicas (O).

Nas sentenças com foco, Cp mostrou-se significativa para diferenciar não só acentuadas (T) de postônicas (O) como também postônicas (O) de pretônicas (E).

A seguir, a FIG. 24 mostra a distribuição dos dados para a variável Op.

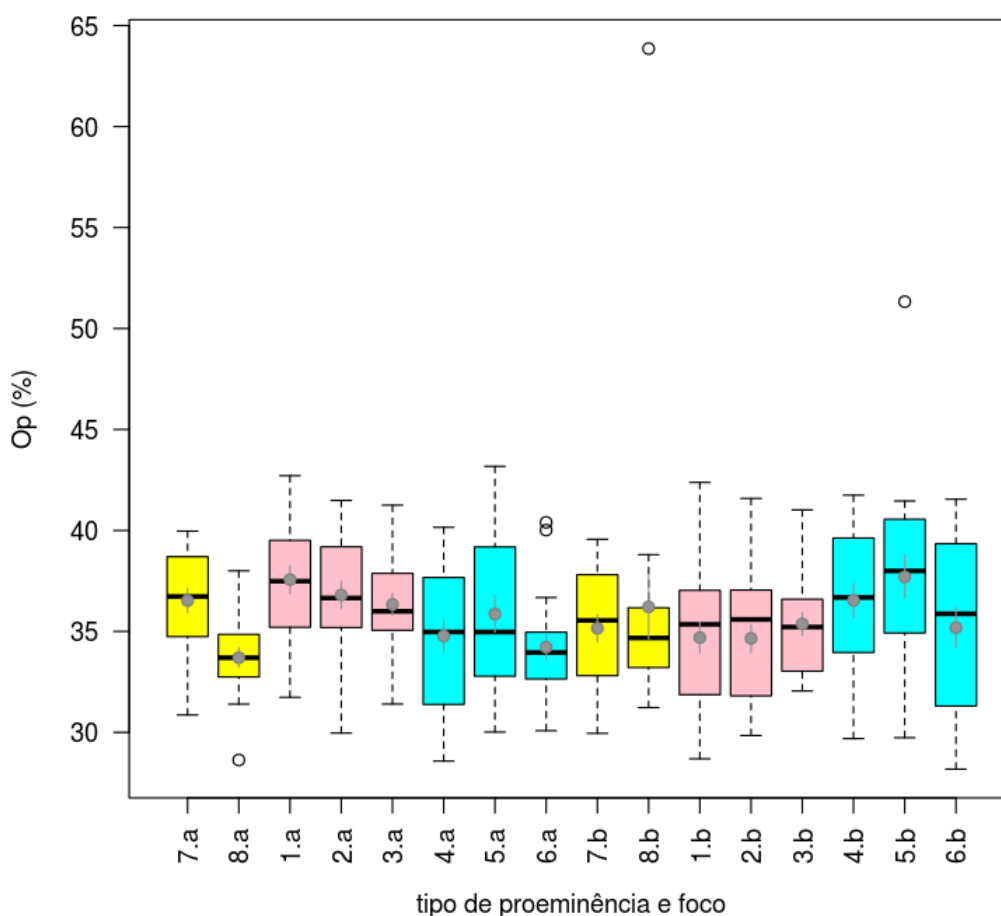


FIGURA 24 - Distribuição dos dados de fase de abertura Op (%) de acordo com tipo de proeminência, localização da proeminência e foco

<i>Tipo de proeminência</i>	<i>Localização da proeminência</i>
acentuada	1 acentuada final
postônica	2 acentuada seguida por uma postônica
pretônica	3 acentuada seguida por duas postônicas
	4 postônica final única
	5 postônica final antecedida por postônica medial
	6 postônica medial
<i>Foco</i>	7 pretônica adjacente à acentuada
a neutra	8 pretônica anterior à sílaba adjacente à acentuada
b com foco	

Na FIG. 24, podemos identificar as seguintes características de distribuição da propriedade fase de abertura (Op): há certa variabilidade em todas as categorias. Verifica-se uma distribuição assimétrica em (2a) exceto nos três casos de pouca variabilidade, mas com *outliers*. Quanto aos centros, aproximam-se as acentuadas neutras seguidas de uma (2a) e duas

postônicas (3a); as postônicas neutras final (4a) e medial (6a); nas sentenças com foco, a acentuada final (1b) e seguida de duas postônicas. A proximidade do centro dessas amostras seria uma pista de que as mesmas não são significativamente diferentes entre si.

Essas observações gerais foram avaliadas por testes simultâneos de comparações múltiplas de médias, que são mostrados na TAB. 9, a seguir.

TABELA 9 - Resultados dos testes de comparações múltiplas para a variável Op

<b>Op</b>						
<b>neutro</b>			<b>foco</b>			
<b>Níveis</b>	<b>valor z</b>	<b>p (&gt; z )</b>		<b>Níveis</b>	<b>valor z</b>	<b>p (&gt; z )</b>
T ~ O	3,860	<1e-04	***	T ~ O	-2,242	0,178
T ~ E	2,338	0,142		T ~ E	-0,850	0,962
O ~ E	-1,938	0,331		O ~ E	1,536	0,607
T1 ~ T2	0,887	0,953		T1 ~ T2	0,033	1,000
T1 ~ T3	1,432	0,682		T1 ~ T3	-0,556	0,996
T2 ~ T3	0,545	0,996		T2 ~ T3	-0,590	0,994
O4 ~ O5	-1,266	0,791		O4 ~ O5	-0,959	0,933
O4 ~ O6	0,539	0,996		O4 ~ O6	1,115	0,872
O5 ~ O6	1,788	0,427		O5 ~ O6	2,074	0,254
E7 ~ E8	-1,266	0,791		E7 ~ E8	-0,959	0,933

Códigos de significância: \*\*\* 0,001 \*\* 0,01 \* 0,05

Como podemos ver na TAB. 9, de modo geral – considerando a significância a  $p \leq 0,05$  –, a variável Cp se mostrou significativa somente para diferenciar acentuadas (T) de postônicas (O).

O QUADRO 6, a seguir, resume os resultados de comparações múltiplas apresentados nas TAB. 3 a 9. As células coloridas indicam as comparações em que a variável dependente foi significativa, ou seja, os tipos de sílabas que a mesma é capaz de distinguir. Células em rosa correspondem às sentenças neutras e células em cinza, às sentenças com foco.

	Dur		F0		Int		CQ		SI		Cp		Op	
	neut	foco	neut	foco	neut	foco	neut	foco	neut	foco	neut	foco	neut	foco
T ~ O	■	■	■	■	■	■	■				■	■	■	
T ~ E	■	■	■			■	■							
O ~ E	■	■	■		■	■						■		
T1 ~ T2		■												
T1 ~ T3			■											
T2 ~ T3			■											
O4 ~ O5				■										
O4 ~ O6														
O5 ~ O6	■	■		■	■									
E7 ~ E8				■										

QUADRO 6 - Resumo dos testes simultâneos de comparações múltiplas de médias

O QUADRO 6 permite identificar um maior número de relações entre as variáveis correspondentes às medidas acústicas (duração, F0, intensidade) e o acento do que entre as variáveis correspondentes às medidas articulatórias (CQ, SI, Cp e Op) e o acento. Consideraremos primeiramente as sentenças neutras. De modo geral, à exceção de SI, todas as variáveis foram significativas para distinguir sílabas acentuadas de sílabas postônicas.

Especificamente com relação às medidas acústicas, vemos que duração e F0 são capazes de distinguir entre os três tipos de sílabas – acentuadas, postônicas e pretônicas –, ao passo que intensidade só distingue acentuadas de postônicas e postônicas de pretônicas – mas não as acentuadas das pretônicas. Duração e intensidade são responsáveis pela distinção entre a postônica final e a medial de uma proparoxítona (O5 ~ O6). Adicionalmente, F0 distingue entre sílabas acentuadas de oxítonas e sílabas acentuadas de proparoxítonas (T1 ~ T3), bem como sílabas acentuadas de paroxítonas e sílabas acentuadas de proparoxítonas (T2 ~ T3).

Quanto às medidas articulatórias, observamos que tanto CQ quanto Cp e Op distinguem entre sílabas acentuadas de sílabas postônicas, mas somente CQ distingue sílabas acentuadas de pretônicas.

Considerando as sentenças com foco, verifica-se que algumas propriedades mudam o papel contrastivo que desempenhavam nas sentenças neutras. Na medidas acústicas, vemos que duração mantém a capacidade de diferenciar acentuadas, postônicas e pretônicas, assim como a postônica final e a medial de uma proparoxítona (O5 ~ O6), mas passa a distinguir

também a sílaba acentuada de uma oxítônica da sílaba acentuada de uma paroxítônica (T1 ~ T2). Intensidade ganha o papel de contrastar as acentuadas das pretônicas, mas perde o de contrastar a postônica final e a medial de uma proparoxítônica (O5 ~ O6). F0 muda quase que radicalmente seu poder distintivo: mantém apenas o papel de contrastar sílabas acentuadas e postônicas e adicionalmente passa a diferenciar a postônica de paroxítonas e a postônica final de proparoxítonas (O4 ~ O5), a postônica final e a medial de uma proparoxítônica (O5 ~ O6), pretônicas adjacente e não adjacente à sílaba acentuada (E7 ~ E8).

Nas medidas articulatórias, vemos que CQ e Op deixam de exercer qualquer papel contrastivo, ao passo que Cp, além de sílabas acentuadas e postônicas, passa a diferenciar também postônicas e pretônicas.

#### 6.2.7 Discussão dos resultados

##### *Medidas acústicas*

A partir dos resultados obtidos neste experimento, verificamos que a propriedade acústica mais significativa na produção do acento no PB, considerando os sujeitos testados, é a duração, corroborando os estudos de Fernandes (1977), Ferreira Neto (2007), Major (1985) e Moraes (1998). Essa é a única propriedade capaz de distinguir entre os três tipos de padrão acentual, tanto em sentenças neutras como em sentenças com foco: sílabas acentuadas são mais longas que as pretônicas e que as postônicas, enquanto as pretônicas são mais longas que as postônicas. Adicionalmente, a duração é capaz de distinguir as sílabas postônicas de proparoxítonas, sendo que as sílabas postônicas finais são mais longas que as mediais.

F0 também é relevante para a distinção dos três padrões acentuais, mas somente em sentenças neutras. Quando é alterada a estrutura prosódica associada à sentença, o que acontece nas sentenças com foco, F0 já não mais distingue os três padrões acentuais, mas passa a contrastar postônicas e pretônicas.

Mostrou-se significativa também a intensidade, mas com menor abrangência: sílabas acentuadas e pretônicas são mais intensas que postônicas. Ao lado da duração, intensidade também diferencia sílabas postônicas de uma proparoxítônica: postônicas finais são mais intensas que as mediais.

Retomaremos agora os resultados de Moraes (1998), apresentados anteriormente nas FIG. 6 e 7 e recuperados a seguir por praticidade.

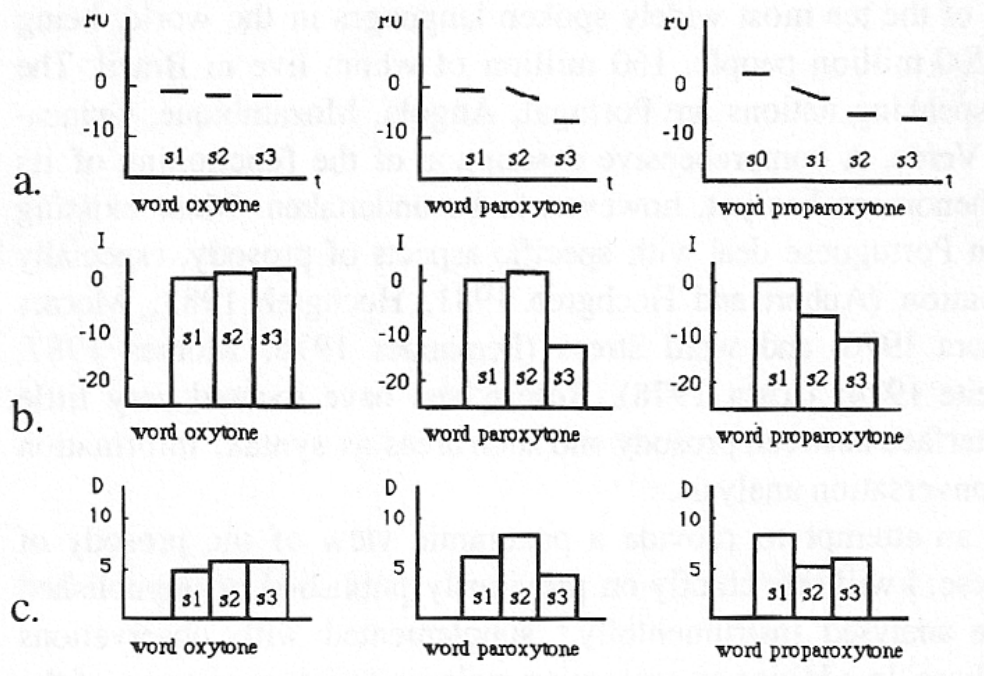


FIGURA 6 (retomada) - Propriedades acústicas do acento lexical em início de sentença  
 a) F0 (quartals de tom); b) Intensidade (dB); c) Duração (cs)

FONTE: MORAES, 1998, p. 180

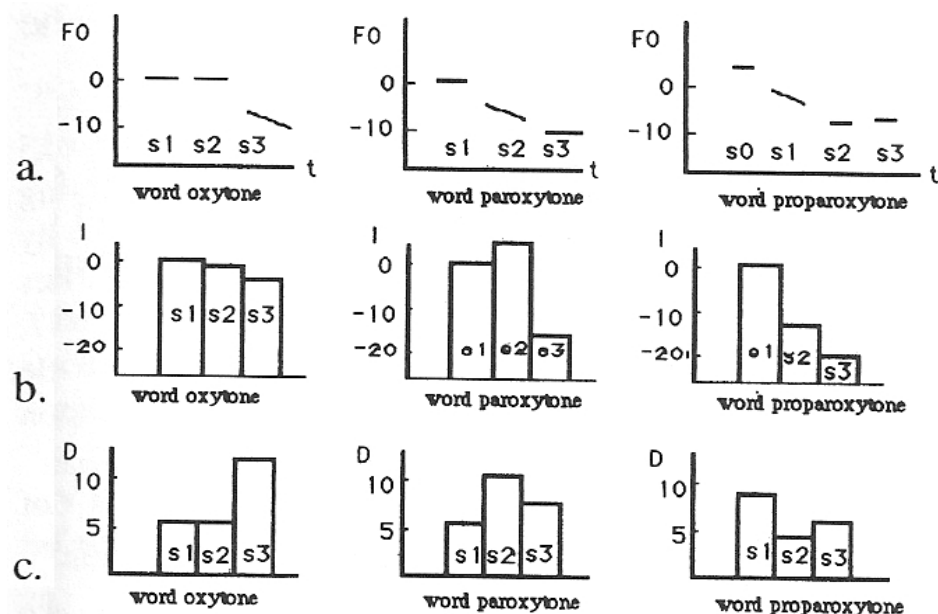


FIGURA 7 (retomada) - Propriedades acústicas do acento lexical em final de sentença  
 a) F0 (quartals de tom); b) Intensidade (dB); c) Duração (cs)

FONTE: MORAES, 1998, p. 18

A partir das medianas das FIG. 18 a 20, elaboramos gráficos semelhantes aos de Moraes, com a finalidade de facilitar a comparação. Tais gráficos são mostrados a seguir, na FIG. 25.

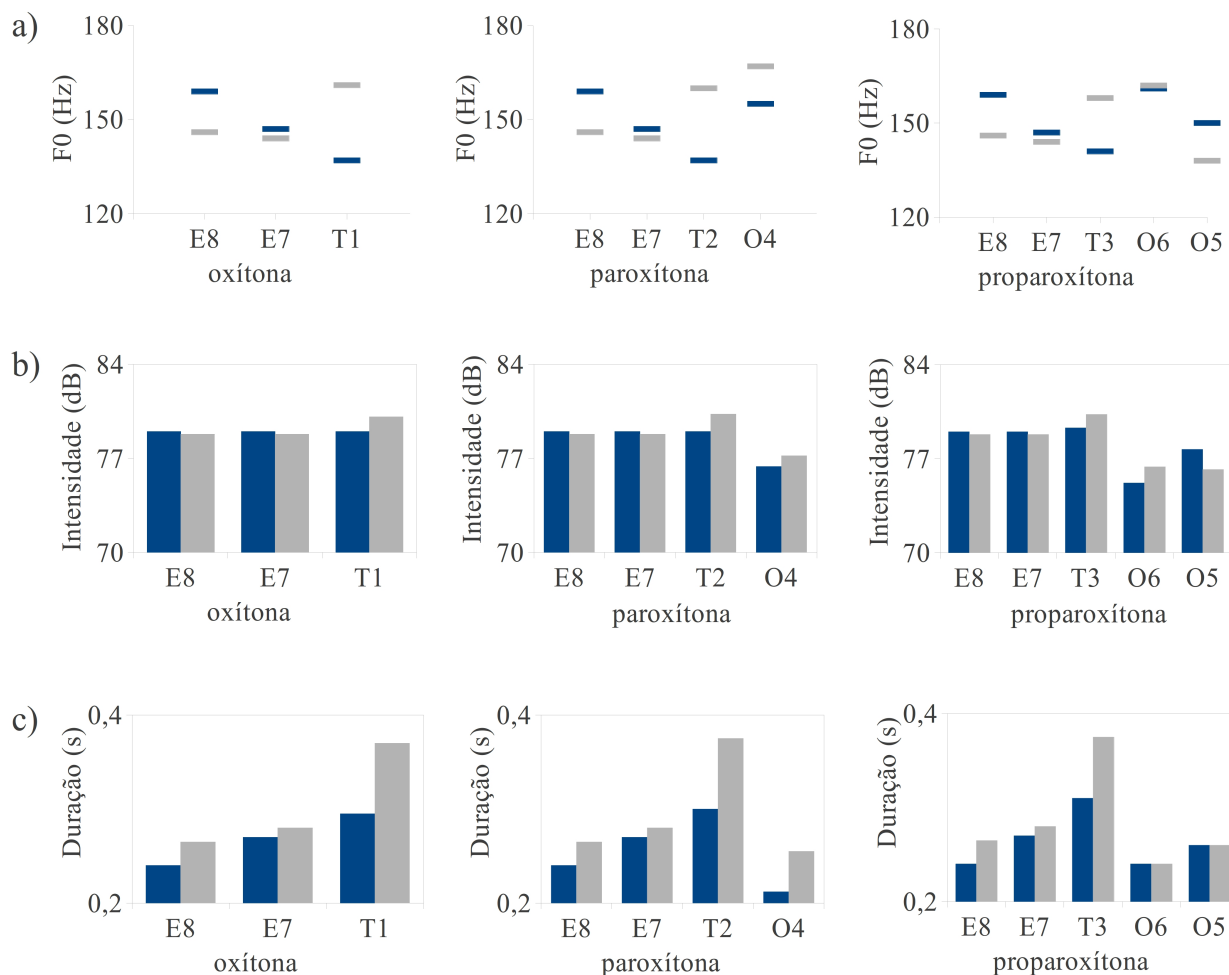


FIGURA 25 - Propriedades acústicas do acento obtidas experimentalmente

a) F0 (Hz); b) Intensidade (dB); c) Duração (s)

*Tipo de proeminência*

T acentuada

O postônica

E pretônica

*Foco*

■ neutra

■ com foco

*Localização da proeminência*

1 acentuada final

2 acentuada seguida por uma postônica

3 acentuada seguida por duas postônicas

4 postônica final única

5 postônica final antecedida por postônica medial

6 postônica medial

7 pretônica adjacente à acentuada

8 pretônica anterior à sílaba adjacente à acentuada

Optamos por unir em um mesmo gráfico os valores correspondentes às sentenças neutras e com foco, ao contrário de Moraes, que separou em gráficos diferentes os padrões

acentuais em início (FIG. 6) e no fim da sentença (FIG. 7). Na FIG. 25, os dados em azul representam as sentenças neutras e os em cinza, as sentenças com foco. Mantivemos os rótulos que viemos adotando para sílabas pretônicas (E), acentuada (T) e postônicas (O):

- a) pretônicas: adjacente à acentuada (E7), antes da adjacente (E8)
- b) acentuadas: em oxítona (T1), em paroxítona (T2), em proparoxítona (T3)
- c) postônicas: em paroxítona (O4), final em proparoxítona (E5), medial em proparoxítona (E6)

Moraes, em seus gráficos, optou por numerar as três últimas sílabas de cada padrão acentual, em ordem crescente da esquerda para direita, como s1, s2, s3. Com isso, mostra duas pretônicas nas oxítonas, uma nas paroxítonas e nenhuma nas proparoxítonas (exceto no gráfico de F0, que inclui uma sílaba extra, anterior à acentuada, nomeada s0). Além disso, deve-se levar em consideração pequenas diferenças nas unidades de medida, especialmente a de F0: adotamos Hz enquanto Moraes adotou quartos de tom. Ainda que não haja uma correspondência perfeita com os rótulos de Moraes, as diferenças não impedem a comparação global de seus resultados com os obtidos no presente estudo. Ademais, iremos assumir a semelhança das propriedades acústicas envolvidas nos aspectos prosódicos comparados por Moraes (início vs. fim de sentença) e os comparados em nosso experimento (neutro vs. foco).

Ao comparar as FIG. 6 e 7, retomadas acima, com a FIG. 25, vemos que há aspectos semelhantes e divergentes entre os resultados. Considerando primeiramente as semelhanças, os dois resultados indicam que, na posição sem ênfase prosódica (início de sentença/neutro):

- a) Duração: postônicas são mais curtas que as pretônicas, e ambas são mais curtas que as acentuadas; postônicas mediais são mais curtas que as finais
- b) Intensidade: postônicas são reduzidas em intensidade com relação a acentuadas e pretônicas
- c) F0: há uma queda em F0 antes da sílaba acentuada

Na posição com ênfase prosódica (fim de sentença/foco), ambos os estudos indicam um aumento significativo dos contrastes entre as sílabas acentuadas e não acentuadas, em decorrência da sobreposição da proeminência entonacional sobre a acentual.

Quanto às diferenças entre os resultados, podemos mencionar, em nossos dados, uma diferença significativa, nas oxítonas, entre as duas sílabas pretônicas e a acentuada, tanto em termos de duração, quanto em termos de F0. Além disso, na situação neutra, a postônica

medial tem menor intensidade que a final, nas proparoxítonas. Esses dois aspectos divergentes podem ser decorrentes de diferenças dialetais, uma vez que os dados de Moraes são da variedade carioca e os nossos, da variedade mineira, que tem como uma das características marcantes de sua fonologia a alta incidência de fenômenos redutivos em sílabas átonas. Os fenômenos redutivos serão melhor explorados mais à frente, quando discutirmos resultados dos experimentos de percepção.

Por fim, observamos em nossos dados que, nas sentenças neutras, há uma curva de F0 que decresce da pretônica para a sílaba acentuada e torna a subir em direção à postônica. Já nas sentenças com foco, a curva tem o mesmo formato, mas sua mudança de trajetória é antecipada, começando a subir em direção à postônica ainda na sílaba acentuada. Os dados de Moraes não mostram o mesmo padrão, talvez porque avaliamos as duas pretônicas mais próximas à sílaba acentuada em todos os padrões acentuais.

### *Medidas articulatórias*

Passaremos, então, a considerar as medidas articulatórias. Retomando as TAB. 3 a 9, que mostram os resultados de testes de comparações múltiplas, observamos que há uma menor relação entre as medidas do EGG e o acento, em comparação com as medidas acústicas. SI (índice de velocidade) não foi significativo em nenhuma circunstância. CQ (quociente de contato) se mostrou significativo, nas sentenças neutras, para diferenciar acentuadas vs. postônicas e acentuadas vs. pretônicas. Cp (fase de fechamento) e Op (fase de abertura) foram significativos nas sentenças neutras, no contraste entre acentuadas vs. postônicas: Cp é menor nas sílabas acentuadas e pretônicas que nas postônicas, enquanto Op é maior no mesmo contexto. Uma vez que SI é uma medida obtida a partir de uma relação entre Cp e Op<sup>31</sup>, a diminuição de um acompanhada pelo aumento do outro pode ter sido um dos fatores que levaram à não significância de SI.

Um de nossos objetivos neste experimento era relacionar propriedades acústicas relevantes na produção do acento a propriedades articulatórias em nível glótico. As propriedades acústicas que identificamos como principais candidatos a uma possível correlação foram intensidade e F0, uma vez que podem ser controladas por ajustes na glote. Contudo, não pudemos observar em nossos resultados uma relação sistemática entre nenhuma

---

31 Retomando:  $SI = (Cp - Op) / (Cp + Op)$

das variáveis articulatórias e as acústicas. CQ é significativo quando intensidade não é (T ~ E) e vice-versa (O ~ E). Sempre que CQ é significativo, F0 também o é, mas o contrário não é verdadeiro. O mesmo acontece com Cp e Op.

O fato de que as medidas articulatórias obtidas a partir do EGG não acompanhem as medidas acústicas talvez possa ser explicado como decorrente dos mecanismos articulatórios e aerodinâmicos empregados para alteração de F0 e intensidade: se essas propriedades acústicas foram modificadas pelos sujeitos assistematicamente, tanto pelo recurso à pressão subglótica, quanto por ajustes nas cordas vocais, podem ter sido mascaradas as consequências de cada mecanismo no sinal do EGG. Infelizmente não é possível testar essa hipótese com os dados de que dispomos no momento, uma vez que não foram realizadas medidas de pressão subglótica.

Iremos finalizar a discussão dos resultados deste experimento avaliando as medidas articulatórias CQ, SI, Cp e Op que obtivemos com os valores de referência adotados (cf. seção anterior). As TAB. 10 a 13, a seguir, apresentam valores médios de cada uma das medidas do EGG, por sujeito e por sílaba, apenas para as sentenças neutras, que são compatíveis com a circunstância de produção dos valores de referência adotados.

TABELA 10 - Valores médios de CQ (%)

Sílaba	Sujeito	T	E	O
<b>GA</b>	<b>E</b>	57,01	51,60	50,88
	<b>G</b>	53,35	52,32	52,72
	<b>J</b>	56,75	51,04	53,76
	<b>R</b>	59,99	54,97	51,10
	<b>V</b>	54,17	52,16	50,14
<b>KU</b>	<b>E</b>	56,59	56,92	55,13
	<b>G</b>	59,37	61,13	60,00
	<b>J</b>	55,95	55,45	51,66
	<b>R</b>	57,56	51,88	49,68
	<b>V</b>	58,54	55,23	57,31
<b>PA</b>	<b>E</b>	57,36	55,90	55,78
	<b>G</b>	51,17	46,64	52,79
	<b>J</b>	56,28	54,88	54,97
	<b>R</b>	59,22	56,97	53,00
	<b>V</b>	54,30	55,49	58,74
<b>PI</b>	<b>E</b>	56,77	55,79	57,57
	<b>G</b>	55,70	51,06	51,73
	<b>J</b>	53,83	54,37	53,85
	<b>R</b>	58,87	52,56	50,93
	<b>V</b>	57,62	55,45	56,07

Como podemos ver na TAB. 10, os valores médios de CQ estão de acordo com o valor de referência ( $\approx 50\%$ ), indo de pouco mais que 45% a pouco mais que 60%. Interessante notar que os valores médios mais altos e mais baixos pertencem aos mesmo sujeito (G). De fato, alguns sujeitos apresentam maior variabilidade nos valores médios de CQ.

TABELA 11 - Valores médios de SI (u.a.)

Sílaba	Sujeito	T	E	O
<b>GA</b>	<b>E</b>	-0,65	-0,65	-0,59
	<b>G</b>	-0,43	-0,36	-0,23
	<b>J</b>	-0,24	-0,28	-0,24
	<b>R</b>	-0,33	-0,30	-0,20
	<b>V</b>	-0,32	-0,37	-0,31
<b>KU</b>	<b>E</b>	-0,63	-0,61	-0,64
	<b>G</b>	-0,27	-0,51	-0,28
	<b>J</b>	-0,35	-0,29	-0,38
	<b>R</b>	-0,27	-0,30	-0,23
	<b>V</b>	-0,41	-0,24	-0,22
<b>PA</b>	<b>E</b>	-0,66	-0,62	-0,71
	<b>G</b>	-0,40	-0,27	-0,25
	<b>J</b>	-0,32	-0,31	-0,31
	<b>R</b>	-0,32	-0,26	-0,14
	<b>V</b>	-0,36	-0,33	-0,40
<b>PI</b>	<b>E</b>	-0,62	-0,68	-0,68
	<b>G</b>	-0,28	-0,12	-0,20
	<b>J</b>	-0,30	-0,33	-0,26
	<b>R</b>	-0,27	-0,35	-0,38
	<b>V</b>	-0,35	-0,45	-0,43

Quanto a SI, podemos ver, na TAB. 11, que os valores médios de quatro dos sujeitos se distanciam do valor de referência ( $\approx -0,65$ ), e giram em torno de  $-0,31$ , sempre negativos. O quinto sujeito (E) apresentou valores que contrastam com os valores dos demais sujeitos e são bem próximos ao valor de referência.

TABELA 12 - Valores médios de Cp (%)

<b>Sílaba</b>	<b>Sujeito</b>	<b>T</b>	<b>E</b>	<b>O</b>
<b>GA</b>	<b>E</b>	8,55	8,12	9,11
	<b>G</b>	16,66	18,66	22,98
	<b>J</b>	21,71	19,10	21,85
	<b>R</b>	19,67	20,57	21,88
	<b>V</b>	19,51	15,59	18,19
<b>KU</b>	<b>E</b>	8,64	9,82	8,00
	<b>G</b>	23,46	22,38	24,15
	<b>J</b>	17,09	19,62	17,99
	<b>R</b>	20,66	18,02	19,80
	<b>V</b>	17,24	21,96	23,80
<b>PA</b>	<b>E</b>	7,97	9,03	6,59
	<b>G</b>	16,71	19,60	23,30
	<b>J</b>	19,07	18,71	20,10
	<b>R</b>	19,74	20,76	24,74
	<b>V</b>	18,41	19,46	17,18
<b>PI</b>	<b>E</b>	9,55	7,55	7,67
	<b>G</b>	21,79	26,01	27,56
	<b>J</b>	19,53	17,34	20,67
	<b>R</b>	20,12	17,82	18,24
	<b>V</b>	20,80	18,45	21,27

A TAB. 12 aponta que, à semelhança do que observamos para SI, os valores de Cp também se mostraram sujeitos à variação individual: os valores médios dos mesmos quatro sujeitos se distanciam do valor de referência (5 a 10%), girando em torno de 21%. O mesmo sujeito (E) apresentou valores que contrastam com os valores dos demais sujeitos e são mais próximos ao valor de referência.

TABELA 13 - Valores médios de Op (%)

<b>Sílaba</b>	<b>Sujeito</b>	<b>T</b>	<b>E</b>	<b>O</b>
<b>GA</b>	<b>E</b>	40,19	35,46	34,17
	<b>G</b>	39,59	35,65	34,91
	<b>J</b>	34,85	32,04	33,93
	<b>R</b>	38,05	37,38	32,27
	<b>V</b>	34,32	34,32	33,07
<b>KU</b>	<b>E</b>	36,86	38,18	37,66
	<b>G</b>	38,75	34,29	33,88
	<b>J</b>	34,71	35,67	32,26
	<b>R</b>	33,72	31,66	30,62
	<b>V</b>	36,84	35,21	35,61
<b>PA</b>	<b>E</b>	37,41	36,24	37,26
	<b>G</b>	35,02	32,85	36,27
	<b>J</b>	36,59	35,50	36,88
	<b>R</b>	36,88	34,92	31,50
	<b>V</b>	36,41	37,63	40,07
<b>PI</b>	<b>E</b>	39,59	39,50	40,16
	<b>G</b>	37,29	32,32	32,51
	<b>J</b>	34,77	34,48	35,07
	<b>R</b>	34,81	36,24	32,70
	<b>V</b>	41,61	37,08	38,77

Na análise da TAB. 13, verificamos que os valores médios de Op giram em torno de 35%, conforme o valor de referência (33 a 39%). Nesse caso, se existem tendências individuais, as mesmas não são evidentes. Contudo, os valores bem distribuídos de Op podem ajudar a explicar tendências individuais que observamos em SI, na TAB. 11: uma vez que SI é calculado a partir da relação entre Cp e Op, pode-se concluir que o contraste dos valores de SI do sujeito E com os demais sujeitos decorre na verdade de uma diferença em Cp, já que os valores de Op são relativamente homogêneos para todos os sujeitos entre si.

Os valores médios de CQ, SI, Cp e Op indicam tendências gerais comuns a todos os sujeitos, em termos da relação dessas variáveis com o acento. Verifica-se também a ocorrência de variação individual e, especialmente, um comportamento divergente de um sujeitos em algumas medidas. Tal comportamento divergente pode ser reflexo de estratégias articulatórias diferentes empregadas na realização de uma mesma propriedade acústica, e pode ter contribuído para que as medidas articulatórias obtidas pelo EGG não se mostrassem tão relacionadas às medidas acústicas como esperado inicialmente.

### 6.2.8 Conclusões

Este experimento teve como objetivo avaliar as propriedades acústicas e articulatórias envolvidas na realização do acento. Procuramos testar a hipótese de que parâmetros acústicos e articulatórios produzidos nas sílabas acentuadas são diferentes das sílabas pretônicas e postônicas. Foram avaliadas três propriedades acústicas (duração; F0; intensidade) e quatro propriedades articulatórias obtidas por meio de eletroglotografia (quociente de contato, CQ; índice de velocidade, SI; fase de fechamento, Cp; fase de abertura, Op). Foram analisadas sílabas acentuadas, postônicas e pretônicas, com e sem ênfase prosódica (foco).

Especificamente, esperava-se que, em comparação com as sílabas átonas, as sílabas acentuadas apresentassem maior duração, maior intensidade e F0 menor nas sentenças neutras e F0 maior nas sentenças com foco. Quanto às medidas específicas do EGG, esperava-se, nas sílabas acentuadas, um maior CQ, menor SI, menor Cp e maior Op, com relação às átonas.

Destacaremos a seguir os aspectos mais importantes dos resultados obtidos neste experimento. Os dados avaliados indicam que a propriedade acústica mais significativa na produção do acento no PB é a duração, como apontado por estudos anteriores sobre a acentuação no PB, como Fernandes (1977), Ferreira Neto (2007), Major (1985), Moraes (1998). Em nossos resultados, duração foi a única propriedade capaz de distinguir entre os três tipos de padrão acentual, tanto em sentenças neutras como em sentenças com foco: sílabas acentuadas são mais longas que as pretônicas e que as postônicas, enquanto as pretônicas são mais longas que as postônicas. Adicionalmente, a duração é capaz de distinguir as sílabas postônicas de proparoxítonas, sendo que as sílabas postônicas finais são mais longas que as mediais.

F0 também se mostrou relevante para a distinção dos três padrões acentuais, mas somente em sentenças neutras. Quando era alterada a estrutura prosódica associada à sentença (i.e., nas sentenças com foco), F0 já não mais distinguia os três padrões acentuais.

Mostrou-se significativa também a intensidade, mas com menor abrangência: sílabas acentuadas e pretônicas são mais intensas que postônicas. Ao lado da duração, intensidade também diferencia sílabas postônicas de uma proparoxítona: postônicas finais são mais intensas que as mediais.

Comparamos os resultados obtidos com os de Moraes (1998) e pudemos observar vários aspectos semelhantes. Na condição sem ênfase prosódica, ambos os estudos indicam que as sílabas acentuadas são mais longas que as pretônicas e estas, mais longas que as postônicas. Entre as postônicas de uma proparoxítona, as finais são mais longas que as mediais. Sílabas acentuadas e pretônicas apresentam maiores valores de intensidade que as postônicas. Há uma queda em F0 antes da sílaba acentuada. Na posição com ênfase prosódica, ambos os estudos indicam um aumento significativo dos contrastes entre as sílabas acentuadas e não acentuadas, em decorrência da sobreposição da proeminência entonacional sobre a acentual. Curiosamente, nas sentenças com foco, verificamos que algumas propriedades acústicas mudaram o papel contrastivo que desempenhavam nas sentenças neutras.

A principal diferença que observamos em nossos resultados com relação aos de Moraes foi a ocorrência de uma curva de F0 decrescente da pretônica até a sílaba acentuada e crescente da acentuada até a postônica, nas sentenças neutras. Nas sentenças com foco, a curva tinha o mesmo formato, mas sua mudança de trajetória era antecipada, começando a subir ainda na sílaba acentuada. Os resultados de Moraes não mostram o mesmo padrão, talvez porque seus dados não contemplam as duas pretônicas mais próximas à sílaba acentuada em todos os padrões acentuais, como avaliamos em nossos dados.

Ao contrário do que era esperado inicialmente, não observamos em nossos resultados uma relação sistemática entre nenhuma das variáveis articulatórias e as propriedades acústicas. Identificamos menos relações entre as variáveis articulatórias e o acento do que entre as variáveis acústicas e o acento.

Constatou-se certa variabilidade individual e valores divergentes de alguns sujeitos nas medidas articulatórias obtidas a partir do EGG. Os valores divergentes foram interpretados como decorrentes de diferentes estratégias articulatórias empregadas para realizar uma mesma propriedade acústica. Sugerimos que essa explicação poderia justificar também o fato de não termos verificado as correspondências diretas esperadas entre propriedades articulatórias e acústicas.

A principal implicação teórica deste experimento para o presente estudo diz respeito ao papel da duração como propriedade acústica mais relevante na produção do acento, papel compartilhado por outras propriedades acústicas, como F0 e intensidade. Segundo Repp (1982), estabelece-se uma relação de troca (*trading relationship*) entre variáveis que desempenham a mesma função contrastiva, de forma que alterações em uma das variáveis

pode ser compensada pelas outras de mesmo papel. Por exemplo, a sensação térmica, que depende tanto da temperatura quanto da umidade relativa do ar, envolveria uma relação de troca entre essas duas variáveis: ao aumentar a temperatura, pode-se abaixar a umidade de forma tal a obter uma mesma sensação térmica. Argumentamos que o acento no PB envolve uma relação de troca entre duração, F0 e intensidade, em um mapeamento de muitos para um.

Da interação entre acento e foco, pudemos verificar que as propriedades acústicas envolvidas na produção do acento estabelecem uma relação complexa em termos da sinalização de proeminência no nível da palavra (acentos) e em nível superior a ela (foco). A proeminência prosódica afeta as propriedades acústicas envolvidas no acento não só aumentando os contrastes já existentes entre os padrões acentuais, mas também alterando o tipo de contraste exercido por cada propriedade acústica. Com isso, uma propriedade acústica (e.g. F0) que em determinada condição prosódica é determinante na diferenciação dois padrões (e.g. acentuadas vs. pretônicas) pode deixar de realizar esse contraste em outra condição prosódica (e.g. foco). F0 oferece um exemplo claro de influência da prosódia na proeminência acentual, uma vez que essa propriedade serve para estabelecer mais de um contraste sonoro no PB (mapeamento um para muitos). Em nossa abordagem, os mapeamentos um para muitos e muitos para um que envolvem as propriedades acústicas duração, intensidade e F0 decorrem da auto-organização do sistema e são viabilizados pela relação de troca estabelecida entre essas propriedades.

De fato, uma das consequências do estabelecimento de uma relação de troca, em um modelo multirrepresentacional, é a possibilidade de que uma propriedade passe a exercer mais de um papel no sistema, em vista da redundância das pistas sonoras. Argumentaremos, a seguir, que a duração, no PB, passou a exercer outros tipos de contrastes, por estar em relação de troca com intensidade e F0.

No cap. 5, apresentamos uma proposta de abordagem multirrepresentacional do acento no português que partia do sistema latino. Nessa proposta, a acentuação do português teria emergido com a auto-organização de variáveis do sistema acentual latino em decorrência de uma série de mudanças fonológicas desencadeadas pela perda da quantidade vocálica. Defendemos que a presença de mais de uma pista acústica para a sílaba acentuada (relação de troca) possibilitou que duração passasse a exercer um papel relevante no contraste entre formas reduzidas por fenômenos de cancelamento final e formas não reduzidas, como apontaremos a seguir.

O latim não tinha oxítonas, que emergiram no português por meio de processos fonológicos de cancelamento final e, especialmente no português brasileiro, por meio do acréscimo lexical de palavras provenientes de línguas indígenas. No PB, fenômenos fonológicos de cancelamento de sílabas átonas continuam em plena atividade, sendo favorecidos por alguns contextos sonoros (cf. os estudos de Souza (2012) e Leite (2006) sobre o cancelamento de vogais pretônicas e postônicas seguidas de sibilantes). Por exemplo, palavras paroxítonas, como *passé*, ['pasi], são transformadas em oxítonas, com a perda da vogal postônica: ['pas]. Surgem potenciais homófonos, como *paz* ['pas] e *passé* ['pas]. Contudo, essa homofonia é apenas aparente. Sabemos de estudos como Souza (2012) e Leite (2006) que a duração opera como detalhe fonético fino no fenômeno de redução vocálica, mantendo a distinção entre as palavras: *paz* ['pas] apresenta vogal com menor duração que a vogal de *passé* ['pa:s] reduzido. A alternância deixa de ser entre presença e ausência de vogal final ou entre oxítone e paroxítone, passando a ser determinada pela duração – de certa forma retomando a propriedade de quantidade vocálica existente anteriormente no latim.

Há consequências da relação de troca também para a percepção, que iremos avaliar nos experimentos específicos descritos na sessão seguinte. Os resultados obtidos neste experimento oferecem contribuições para a compreensão dos correlatos acústicos e articulatórios do acento no PB e sua interação com aspectos entonacionais e outros fenômenos sonoros da língua.

### 6.3 Experimentos de percepção

Por meio de uma série de três experimentos de percepção, este estudo investigou os mecanismos de processamento e categorização do acento lexical no português brasileiro. Mais especificamente, buscou-se avaliar possíveis influências da estrutura linguística na percepção de proeminência em padrões sonoros não linguísticos (experimento 1) e linguísticos (experimentos 2 e 3). Os dois primeiros experimentos se baseiam na desambiguação de sequências sonoras ambíguas, enquanto o terceiro, na identificação de um intervalo de estímulos progressivamente distintos. Os três experimentos utilizaram um mesmo projeto experimental, representado pela FIG. 26, a seguir.

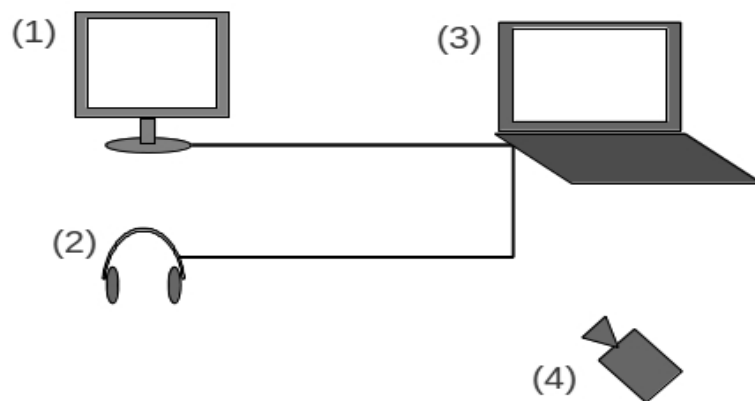


FIGURA 26 - Projeto experimental utilizado nos três experimentos de percepção

Os sujeitos se posicionavam em frente ao monitor (1), que apresentava os estímulos visuais, portando os fones de ouvido (2), pelos quais escutavam os estímulos auditivos. Os experimentos eram controlados pelo computador (3), com extensão de tela no monitor (1), e gravados em vídeo, por meio de câmera digital (4), para documentação. Uma pequena alteração na montagem foi introduzida no terceiro experimento, que continha tarefa auto-monitorada (*self-pacing*), como veremos adiante. Neste caso, foi retirado o monitor (1) em tela estendida; os sujeitos receberam os estímulos visuais e auto-controlaram o experimento diretamente do computador (3).

### 6.3.1 Experimento 1: desambiguação de padrões não linguísticos

O primeiro experimento de percepção foi projetado de forma a oferecer resultados comparáveis aos de Bell (1977), que estudou possíveis influências da estrutura linguística na interpretação prosódica de estímulos não linguísticos. O autor testou a percepção de combinações simples de tons puros que simulavam a proeminência acentual em grupos de falantes de cinco línguas diferentes. Como o português não se encontrava entre as sete línguas testadas pelo autor, um de nossos objetivos com o presente experimento é avaliar o comportamento dessa língua entre os resultados do estudo original.

Estruturas rítmicas são formadas pela repetição de sequências de sons em que alternam elementos semelhantes e diferentes. Sabemos de estudos em psicoacústica que sequências de sons tendem a ser quebradas (ou *parsed*) em agrupamentos menores, em que um dos elementos geralmente é mais saliente (HANDEL, 1993). Como relata o autor, os agrupamentos tendem a ser compostos por dois, três ou quatro elementos. Um agrupamento binário é formado de dois elementos diferentes, geralmente sendo a um deles conferida maior saliência cognitiva. O elemento mais saliente pode se posicionar no início (i.e. *sw*<sup>32</sup>) ou no fim da sequência (i.e. *ws*). Por sua vez, um agrupamento ternário é composto de dois elementos iguais e um diferente, podendo este último, o mais saliente, se localizar no início (i.e. *sww*), meio (i.e. *wsw*) ou fim da sequência (i.e. *wws*). Esses dois tipos de agrupamentos, binário e ternário, são os de maior interesse para o presente estudo, uma vez que guardam correspondência com os padrões acentuais mais frequentes no PB.

Bell (1977) identificou em seus resultados uma preferência geral para a acentuação inicial dos grupos na sequência controle, em que nada variava, assim como nas sequências em que intensidade e frequência eram alteradas. Por outro lado, houve uma tendência generalizada à acentuação medial dos grupos na sequência em que a duração variava. Analisando os resultados obtidos, o autor afirma que “the tendencies that appear for each language do not follow the pattern of the language’s accent placement” e concluiu que “the effect of accentual structure upon rhythmic perception of non-speech sounds, if it exists, is unlikely to be of the same magnitude as the general psychophysical effects discussed earlier.”

---

32 Nos padrões sonoros, a relação de saliência cognitiva será marcada por meio da notação “s”, para saliente (“*strong*”) e “w”, para não (ou menos) saliente (“*weak*”), ainda que não esteja em jogo necessariamente uma relação de força.

(BELL, 1977, p. 7). Por outro lado, Jakobson, Fant e Halle (1952), mencionados por Bell (1977), defendem que o sistema acentual da língua de um falante exerce influências sobre a percepção de padrões rítmicos em geral, mesmo os não linguísticos. Considerando os modelos teóricos adotados neste estudo, para os quais as representações são multirrepresentacionais e a linguagem se baseia em processos cognitivos gerais, espera-se que a percepção não linguística esteja sujeita a influências da estrutura linguística.

Este primeiro experimento de percepção pretende testar as seguintes hipóteses: (1) a percepção de padrões sonoros é afetada diferentemente por alterações nas propriedades acústicas; (2) em padrões não linguísticos contínuos do tipo ternário, há uma maior probabilidade de percepção do padrão com proeminência medial (*ws*, i.e. *weak-strong-weak*), correspondente ao padrão paroxítono, que apresenta maior frequência de tipo no PB. Esta segunda hipótese equivaleria a testar se padrões rítmicos da língua influenciam a percepção rítmica (linguística e não linguística), como defendido por Jakobson, Fant e Halle (1952). No contexto dos modelos teóricos adotados neste estudo – em especial, os Modelos de Uso (BARLOW; KEMMER, 2000; BYBEE, 2001), para os quais a frequência de uso desempenha um papel crucial na construção da gramática do falante –, o presente experimento tem potencialmente como interesse adicional avaliar efeitos de frequência de tipo sem influência da frequência de ocorrência, uma vez que os estímulos não são linguísticos.

#### 6.3.1.1 Sujeitos

Foram testados 12 sujeitos, sendo 6 homens e 6 mulheres, todos com nível universitário, na faixa etária entre 20 e 30 anos. Todos são falantes nativos do português e não relataram histórico de problemas de audição ou fala. Além disso, todos os sujeitos são leigos em teoria e instrumentos musicais. Este fator foi controlado, em vista de possíveis diferenças no desempenho de músicos em comparação a não músicos em tarefas envolvendo estímulos não linguísticos. Os mesmos sujeitos participaram deste experimento e do experimento 2, em uma mesma sessão de coleta.

### 6.3.1.2 Estímulos

Os estímulos consistiram em oito sequências de tons puros, sendo uma sequência controle (em que nada variava entre os elementos do grupo) e sete sequências testes, nas quais um tom a cada três variava, simultaneamente ou não, em duração, frequência e/ou intensidade. A FIG. 27, abaixo, ilustra esquematicamente cada sequência de tons construída.

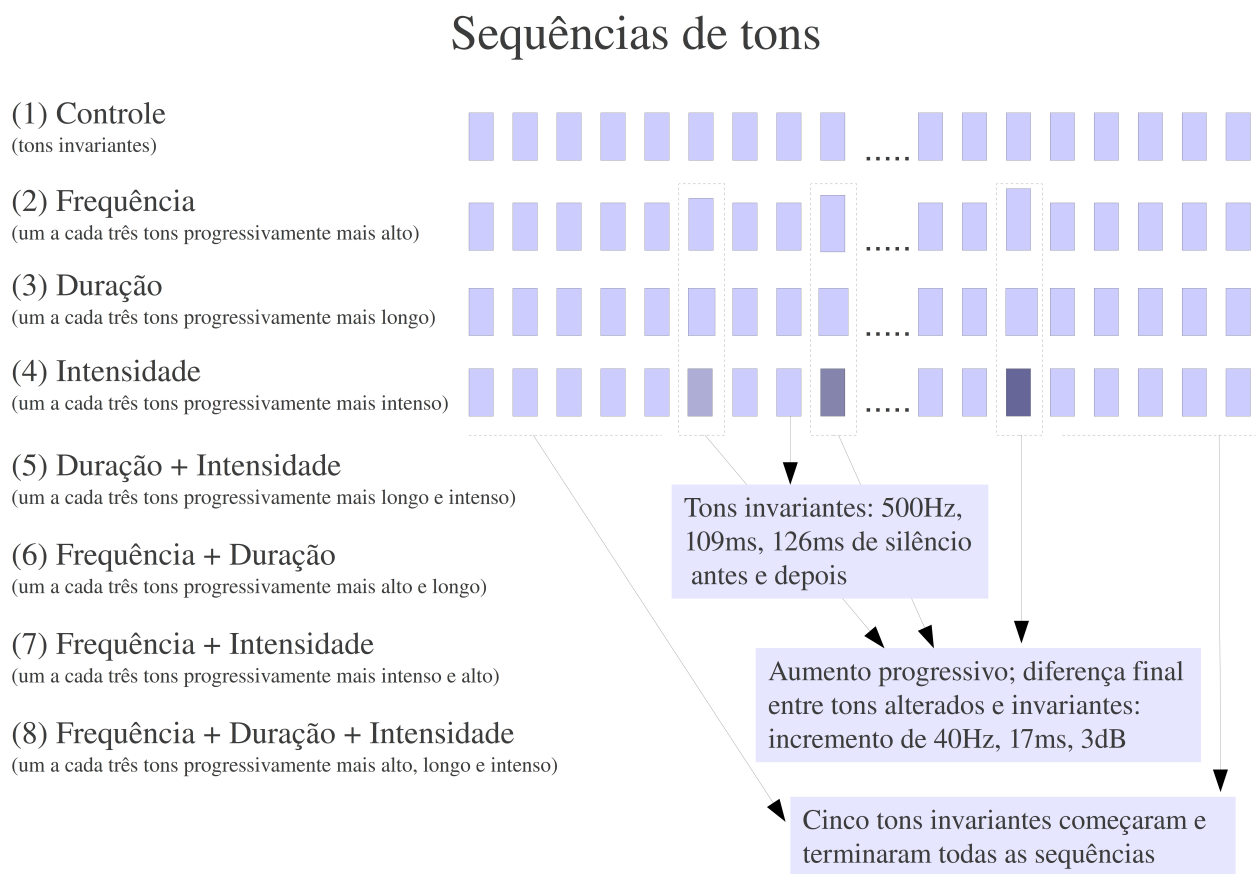


FIGURA 27 - Sequências de tons utilizadas como estímulos no experimento de percepção 1

### 6.3.1.3 Procedimentos

O teste realizado no presente experimento segue a metodologia de Bell, com algumas adaptações. A tarefa consistiu em escutar uma sequência de tons ambígua em termos de *parsing* da proeminência (e.g. ...wwswswswsw... → ...[wws][wws][wws]ww... ou ...w[ws] [ws]w[ws]w... ou ...ww[sw][sw][sw]...) e em seguida realizar a decisão forçada por um

entre quatro padrões apresentados, nos quais variava o elemento mais proeminente. Neste experimento, os sujeitos levaram cerca de 15 minutos para completar a tarefa solicitada.

Os sujeitos escutavam, por meio de fones de ouvido, sete tipos de sequências de três tons – como mostra a FIG. 27, acima –, que variavam sistematicamente em F0, intensidade e/ou duração. Cada sequência foi formada por dois tons invariáveis e um terceiro tom, variável, alterado em um ou mais parâmetros acústicos. Entre cada tom havia um intervalo de silêncio de 126ms nos tons invariantes, sendo reduzido apenas nos tons em que foi alterada a duração, de forma a manter constante, em todas as sequências, a soma da duração dos tons com os intervalos entre eles, como fez Bell (1977). Quando tocadas sucessivamente, as sequências apresentam ambiguidade em sua segmentação com relação à localização do tom variável perante os dois invariantes. Após escutar cada sequência de tons, os sujeitos deviam categorizar cada sequência ouvida, apontando para um dos quatro diagramas apresentados simultaneamente na tela de um computador. Tais diagramas correspondiam à estrutura xxx; wws; wsw; sww (sendo x = tom invariável; s = tom saliente; w = tom não saliente), como mostra a FIG. 28, na sequência.



FIGURA 28 - Estímulos visuais utilizados  
no experimento de percepção 1

#### 6.3.1.4 Resultados

Neste experimento, a variável dependente é a resposta ao padrão sonoro como neutro (XXX), proeminência final (WWS), proeminência medial (WSW), ou proeminência inicial

(SWW). A variável independente é a natureza acústica da proeminência do padrão sonoro, que poderia ser realizada apenas por frequência (F), duração (D) ou intensidade (I), isoladamente, ou por meio das interações entre os três níveis: FD, FI, DI, FDI. A variável “natureza acústica” contém ainda um último nível, referente ao padrão sonoro controle (C), em que não havia relação de proeminência.

Os dados foram ajustados por um modelo generalizado linear com três fatores fixos<sup>33</sup>, representado em (i). Como se pode ver, os níveis da variável independente “natureza acústica” foram inseridos individualmente no modelo, transformados em fatores fixos.

(i) resposta ~ frequência + duração + intensidade

Em decorrência de restrições do pacote de funções utilizado para análise dos dados, que não permite a incorporação de fatores aleatórios, o modelo contém apenas os três fatores fixos frequência, duração e intensidade. Ou seja, os informantes não puderam ser considerados como fator aleatório no modelo, o que poderia ter prevenido uma possível inflação dos graus de liberdade dos resíduos.

Adotou-se um modelo generalizado porque a variável dependente são proporções de escolha de cada padrão sonoro, com distribuição de erros multinomial: uma vez que as respostas poderiam ser de quatro tipos diferentes, a variável dependente é um vetor de proporções das três probabilidades de resposta referentes aos padrões acentuais: P(WWS), P(WSW) e P(SWW). Obtém-se a probabilidade de resposta referente ao padrão neutro a partir das outras três probabilidades:  $P(XXX) = 1 - P(WWS) - P(WSW) - P(SWW)$ .

Antes, contudo, de ajustar os dados ao modelo linear, foi necessário transformar os valores de probabilidades, que se restringem ao intervalo [0, 1], a um intervalo  $[-\infty, \infty]$ . Para tal, foi realizada a seguinte transformação logarítmica:

$$\log[P(\text{padrão})/P(XXX)]$$

Após a transformação dos dados, o modelo linear foi ajustado, criando-se três vetores, um para F, um para D, outro para I, assim como um vetor intercepto, que corresponde a (C). Um dos benefícios do modelo (i) adotado é possibilitar a previsão do resultado dos padrões sonoros formados pela interação de efeitos a partir da soma dos efeitos individuais de F, D e I. É possível obter, por exemplo, o efeito de FD (frequência-duração) somando os efeitos de F e D. Por requerer um número menor de coeficientes, é preferível ao modelo (ii),

<sup>33</sup> R: função *vglm*, pacote *VGAM*

que realiza o ajuste para cada um dos coeficientes das interações.

(ii) resposta  $\sim$  frequência \* duração \* intensidade

Os modelos (i) e (ii) foram comparados por meio do teste de razão de verossimilhança<sup>34</sup>, que respeita a distribuição de qui-quadrado. O valor  $p$  obtido é  $> 0.813$ , o que indica que os modelos não são significativamente diferentes entre si e justifica, portanto, a utilização do modelo mais econômico (i).

A FIG. 29, a seguir, mostra os resultados do ajuste ao modelo, na forma de um gráfico triangular, que representa a distribuição das proporções previstas para os efeitos do modelo e suas combinações.

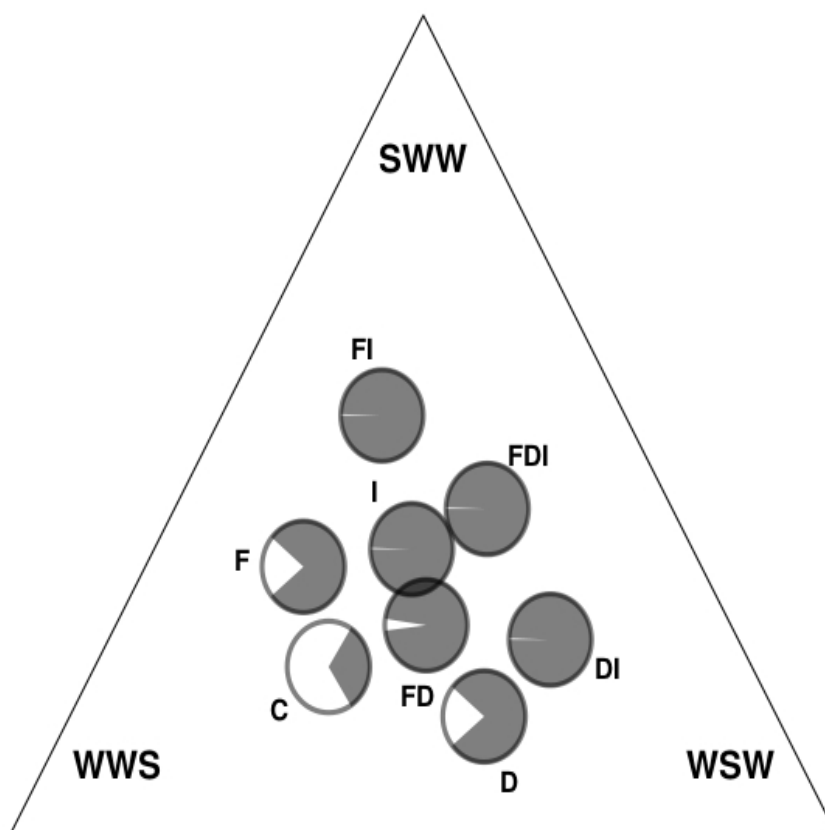


FIGURA 29 - Distribuição das probabilidades de resposta de acordo com o padrão sonoro dos estímulos

F = frequência; D = duração; I = intensidade

Equivalências com padrões acentuais: WWS = oxítone;

WSW = paroxítone; SWW = proparoxítone

<sup>34</sup> R: função *pchisq*, pacote *stats*

Os vértices do gráfico correspondem às três possibilidades de resposta (SSW, WSW ou WWS) referentes à variável dependente. Há sete pontos dentro do triângulo que correspondem aos níveis da variável independente, referente à natureza da proeminência dos padrões sonoros (F, D, I, FD, FI, DI, FDI). O oitavo ponto corresponde ao intercepto (C), referente ao estímulo controle, em que não havia variação. Os pontos são posicionados no interior do triângulo de acordo com a probabilidade de escolha das respostas: a distância entre um ponto e o lado oposto a um vértice é proporcional a  $1 - P$ , sendo  $P$  a proporção de escolha de tal vértice. Ou seja, a probabilidade de que um estímulo seja reconhecido como um dos padrões é proporcional a sua proximidade ao vértice desse padrão. Dessa forma, é possível avaliar as diferenças entre as variáveis em termos de sua distância com relação ao ponto C e em termos da direção do vetor formado entre elas e C em relação aos vértices. Por fim, a cada ponto é sobreposto um “subgráfico” de pizza, cuja porção branca representa a probabilidade de escolha da resposta neutra (XXX).

Algumas tendências podem ser observadas na FIG. 29. Exploraremos as características dos vetores formados por cada variável e o intercepto (C). O vetor  $C \rightarrow D$  se direciona para o vértice WSW. Isto indica que a mudança da condição Controle © à condição Duração (D) favorece o padrão WSW, com proeminência medial, correspondente às paroxítonas. Já o vetor  $C \rightarrow F$  tem como direção  $WSW \rightarrow SWW$ . Ou seja, com relação à condição Controle (C), a condição Frequência (F) favorece as respostas SWW e desfavorece as respostas WSW. O vetor  $C \rightarrow I$ , por sua vez, aponta em direção oposta ao vértice WWS. Isto significa que a condição Intensidade (I) favorece as respostas WWS.

O modelo ajustado não contém interações entre as variáveis básicas (F, D, I), de forma que os padrões conjugados (FI, FD, DI e FDI) podem ser previstos em função dos três vetores de base:  $C \rightarrow F$ ,  $C \rightarrow D$  e  $C \rightarrow I$ .

Uma última observação a respeito da distribuição dos dados pode ser feita. Considerando as respostas ao padrão controle (C) diferentes de XXX, vemos que o mesmo se coloca mais próximo de WWS que das demais respostas.

### 6.3.1.5 Discussão dos resultados

Neste experimento, procurou-se testar as hipóteses de que (1) a percepção de padrões sonoros é afetada diferentemente por alterações nas propriedades acústicas; (2) padrões rítmicos da língua influenciam a percepção rítmica (linguística e não linguística). Enquanto a primeira hipótese pode ser testada diretamente, a segunda deve levar em consideração o sistema acentual do PB, cujas propriedades acústicas e articulatórias foram abordadas neste estudo por meio de experimento de produção. Sabemos da coleta de dados realizada junto ao ASPA (cf. sessão 5.2) que o padrão acentual paroxítono é o mais frequente no PB, representando 71,4% dos verbos e 67,1% dos não verbos. Portanto, a segunda hipótese equivale encontrar uma maior probabilidade de percepção do padrão com proeminência medial (WSW), correspondente ao padrão paroxítono. Essa hipótese vai ao encontro do que defendem Jakobson, Fant e Halle (1952) e contraria os resultados apresentados por Bell (1977).

Os resultados obtidos neste experimento indicam que as propriedades acústicas exercem influência na percepção da proeminência em padrões sonoros, em termos de tendências gerais. Observamos que, isoladamente, a frequência desfavorece a percepção do estímulo como proeminência medial, tendendo a favorecer a proeminência final. A duração favorece a percepção como proeminência medial. A intensidade desfavorece a percepção como proeminência final. O padrão controle, uma sequência sonora de sons idênticos, favorece a percepção de ausência de proeminência em mais de 60% das respostas, como esperado, mas, nos demais casos, favorece a percepção da proeminência final. O QUADRO 7, a seguir, resume as interpretações das relações entre estímulo e proeminência percebida apresentadas acima e as previsões para interações entre fatores.

	<b>SWW</b>	<b>WSW</b>	<b>WWS</b>
<b>F</b>	favorece	desfavorece	
<b>D</b>		favorece	
<b>I</b>			desfavorece
<b>FI</b>	favorece		
<b>FD</b>			desfavorece
<b>DI</b>		favorece	desfavorece
<b>FDI</b>			desfavorece

QUADRO 7 - Tendências de favorecimento e desfavorecimento à percepção padrões

F = frequência; D = duração; I = intensidade;

FI = frequência+intensidade; FD = frequência+duração;

DI = duração+intensidade; FDI = frequência+duração+intensidade

O QUADRO 7 indica que D e DI favorecem WSW, ao passo que F e FI favorecem SWW. Além disso, I, FD, DI, FDI desfavorecem WWS. Podemos apontar uma tendência de que, quando está em jogo a duração, os sujeitos tendem a perceber a proeminência medial, que corresponde ao padrão paroxítono, em detrimento da proeminência final, que corresponde à oxítone. Por outro lado, quando estava em jogo a frequência, os sujeitos tendiam a perceber a proeminência inicial, correspondente à proparoxítone.

Retomando o estudo de Bell (1977), cuja metodologia foi adotada no presente experimento, há algumas diferenças entre nossos resultados e os obtidos pelo autor. Primeiramente, cabe lembrar que o autor não avaliou estímulos em que mais de uma propriedade variava, como fizemos. Bell identificou em seus resultados uma preferência geral de associar C à proeminência inicial, o que não aconteceu em nosso experimento – visto que na FIG. 29 o ponto correspondente a Controle encontra-se mais próximo do vértice WWS, havendo, portanto, maior probabilidade de ser identificado como proeminência final. O autor verificou uma preferência pela proeminência inicial também para F e I, o que curiosamente não aconteceu em nossos estímulos equivalentes, com essas propriedades isoladas, mas aconteceu quando elas eram conjugadas, em FI. Por outro lado, Bell observou uma tendência generalizada à proeminência medial quando a duração variava, o que pode ser observado em nossos dados, estendendo-se ainda ao par conjugado DI.

Em uma análise geral, nossos resultados não evidenciam uma correspondência clara

e direta entre os padrões de proeminência preferentemente percebidos em cada sequência sonora e os padrões acentuais do português, mas aponta tendências gerais condizentes com o sistema acentual dessa língua: as paroxítonas são o padrão mais frequente de acento no domínio da palavra, que tem como principal propriedade acústica a duração. Ressalte-se que, em tons puros, a percepção conta prioritariamente com propriedades físicas, ao passo que, nos eventos linguísticos, o significado é um fator determinante para a identificação dos padrões sonoros, também influenciada pelo uso linguístico. Portanto, no caso de experimentos de percepção envolvendo estímulos não linguísticos, mesmo que haja influência de estruturas específicas da língua, não se deve esperar relações diretas. É necessário considerar a força antagônica exercida pela inseparabilidade do elo arbitrário entre padrões sonoros e o significado, que tem consequências para a percepção: na ausência de significado linguístico, há uma tendência de que os padrões sonoros sejam interpretados como eventos não linguísticos – portanto, independentes da estrutura linguística. Sugerimos que a interação entre a força exercida pela estrutura linguística e a força exercida pela interpretação não linguística tem como consequência a emergência de padrões complexos de resposta. Ao lado de questões metodológicas, como o número reduzido de sujeitos, esse que pode ser um dos fatores responsáveis pelo resultado que Bell (1977) interpretou como evidência contrária à influência da linguagem na percepção de padrões rítmicos não linguísticos.

Em uma análise caso a caso, podemos identificar correspondências entre as tendências na percepção de padrões e a estrutura linguística do português: como sabemos do experimento de produção, apresentado na seção anterior, a duração é a propriedade acústica mais relevante na produção da proeminência acentual no português, que tem como padrão acentual mais frequente as paroxítonas. No presente experimento, houve uma tendência à percepção da proeminência medial, equivalente a paroxítonas, quando avaliamos a duração isoladamente ou conjugada com intensidade. Como menciona Cutler (2005), um resultado aproximado foi obtido por Turk e Sawusch (1996), que apontam para a duração ou a associação de amplitude e duração como propriedades relevantes em julgamentos de proeminência. Do mesmo modo, F<sub>0</sub>, que exerceria funções variadas no sistema sonoro do PB, parece alternar ora entre o favor das oxítonas, ora em favor das proparoxítonas.

### 6.3.2 Experimento 2: desambiguação de padrões linguísticos

O segundo experimento de percepção teve como propósito avaliar a percepção da proeminência em sequências sonoras linguísticas. O paradigma empregado é semelhante ao do primeiro experimento, mas trocando os tons puros por estímulos linguísticos. Com isso, além da frequência de tipo, torna-se possível acessar diretamente também possíveis influências da frequência de ocorrência na percepção de padrões sonoros. Pretende-se testar a seguinte hipótese: nos padrões linguísticos acentuais, a estrutura percebida sofre influência da frequência de ocorrência e de tipo das palavras testadas.

#### 6.3.2.1 Sujeitos

Foram testados os mesmos sujeitos do primeiro experimento de percepção, na mesma sessão de coleta de dados: 12 sujeitos, sendo 6 homens e 6 mulheres, todos com nível universitário, na faixa etária entre 20 e 30 anos. Todos são falantes nativos do português e não relataram histórico de problemas de audição ou fala.

#### 6.3.2.2 Estímulos

Os estímulos utilizados neste experimento seguiram o mesmo princípio do experimento anterior, que avaliou a percepção de sequências de tons puros ambíguas em termos de *parsing* da proeminência, mas, no presente caso, eram sequências de sons linguísticos, interpretáveis como palavras (e.g. ...pɔsipɔsipɔsipɔsipɔs... → ..pɔ[sipɔ][sipɔ][sipɔ][sipɔ]s... “cipó” ou ..[pɔsi][pɔsi][pɔsi][pɔsi]pɔs... “posse”). Foram selecionados como estímulos dez sequências de sílabas fracas e fortes com segmentação ambígua, ou seja, que pudessem ser interpretadas como dois padrões acentuais diferentes, e, conseqüentemente, como duas palavras distintas, a depender do *parsing* realizado no momento da percepção (e.g. ...sipɔsipɔsipɔsi... → “cipó”, oxítone, ou “posse”, paroxítone). Portanto, a cada estímulo,

estavam associadas duas palavras quanto à possibilidade de identificação: uma oxítone e outra paroxítone. Considerando que a frequência de ocorrência poderia influenciar a identificação entre uma das duas palavras, foram selecionados pares de palavras com valores de ocorrência tanto próximos quanto muito distantes, de modo a avaliar se haveria influência desse fator na identificação das palavras pelos informantes. A frequência de ocorrência das palavras, obtida a partir do banco de dado do Projeto Aspa, já mencionado anteriormente (cf. cap. 4), é mostrada no QUADRO 8, a seguir, para cada uma das palavras associadas a cada estímulo. Como se vê na coluna “Estímulo”, foi adotado, para cada estímulo, um código de identificação que corresponde às três primeiras letras da palavra oxítone associada a tal estímulo.

<b>Estímulo</b>	<b>Oxítone</b>	<b>Frequência de Ocorrência</b>	<b>Paroxítone</b>	<b>Frequência de Ocorrência</b>
man	manoel	7.958	noelma	16
dav	davi	1.373	vida	112.365
caj	caju	873	juca	1.922
tum	tumor	2.212	morto	20.475
dal	dali	2.161	lida	1.775
tup	tupi	1.420	pito	206
bab	babá	1.283	baba	432
mar	marrom	1.111	roma	9.689
mac	macau	878	calma	3.501
jab	jabuti	422	botija	43
fub	fubá	357	bafo	205
tut	tutor	312	torto	783
cip	cipó	271	posse	15.246
gui	guichê	227	chegue	2.470
cas	cassis	125	cisca	10
sum	sumô	127	moço	1.639
caq	caqui	107	quica	22
mas	maçom	20	soma	7.913
cja	cajá	59	jaca	127

QUADRO 8 - Pares de estímulos utilizados no experimento de percepção 2 e seus valores de ocorrência

Para construir cada estímulo, as sílabas que compunham as palavras ambíguas foram gravadas isoladamente e, em seguida, concatenadas. A diferença entre a sílaba forte e a sílaba

fraca foi posteriormente reconstruída por meio de manipulação do sinal, utilizando como referência valores de intensidade e duração, respectivamente, das sílabas forte e fraca de uma palavra gravada separadamente. A locutora da gravação foi a própria pesquisadora – 27 anos, falante nativa da variedade de Belo Horizonte. Optou-se por gravar as sílabas isoladamente, e não as próprias palavras, a fim de evitar que possíveis detalhes fonéticos finos característicos de cada palavra e decorrentes da ordenação das sílabas (e.g. transições de formantes) pudessem influenciar na tarefa de identificação dos estímulos. Uma vez gravadas as sílabas, foram realizados os seguintes procedimentos de manipulação, de maneira uniforme para todos os estímulos, por meio do *software* Audacity (AUDACITY Team, 2008), com objetivo de simular a relação de proeminência para cada estímulo, a sílaba forte da sequência foi amplificada em 3dB e a sílaba fraca teve, então, sua amplitude escalonada em 0,5 com relação à forte. Além disso, a duração das sílabas foi alterada por meio da função “Alterar tempo sem alterar tom”: as sílabas fortes passaram a ter 0,4 s e as fracas, 0,3 s. As sílabas forte e fraca foram concatenadas e repetidas 25 vezes cada uma. Por fim, duas medidas foram tomadas para mascarar o início e o fim das sequências, de forma a evitar que os sujeitos fossem influenciados pela ordem inicial ou final para identificação do estímulo, como relatado no experimento de Bell (1977). Primeiro, o início e o fim das sequências foram submetidos respectivamente a efeitos de *fade-in* (atenuação regressiva do sinal) e *fade-out* (atenuação progressiva do sinal). Segundo, a sequência foi transformada de mono para stereo e, no canal adicional, foram introduzidos 3 s de ruído browniano alinhados com o início e com o fim das sequências, tendo, respectivamente, seu final submetido a *fade-out* e início submetido a *fade-in*. A FIG. 30, a seguir, apresenta o sinal construído para um dos estímulos utilizados no experimento 2.

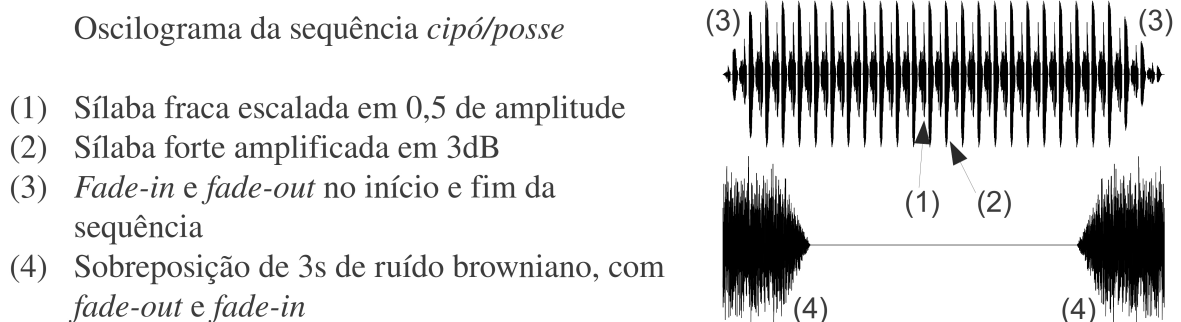


FIGURA 30 - Exemplo de estímulo utilizado no experimento de percepção 2

### 6.3.2.3 Procedimentos

Este experimento consistiu em um teste de identificação, em que os sujeitos escutavam estímulos auditivos e deveriam relatar que palavra escutaram, em tarefa de livre seleção. Os sujeitos escutavam, por meio de fones de ouvido, cada uma das sequências sonoras ambíguas. Em seguida, deviam identificar e relatar a palavra ouvida, e a resposta era anotada pela pesquisadora. Neste experimento, assim como no experimento de percepção 1, os sujeitos levaram cerca de 15 minutos para completar a tarefa solicitada.

Como a tarefa proposta era de livre seleção, foram obtidas respostas que não corresponderam a nenhuma das duas palavras associadas ao estímulo: um total de 34 casos de não correspondência, que não foram consideradas na análise dos dados. Para avaliar o desempenho da tarefa, foi calculada uma taxa de identificação por sujeito e por palavra. A taxa de identificação foi determinada como a porcentagem de respostas do sujeito ou da palavra que correspondessem a uma das duas palavras esperadas. Dois sujeitos apresentaram taxa de identificação próxima de 50% – o que significa que quase a metade de suas respostas não foram aproveitadas. Juntos, responderam por 17 do total de 34 palavras sem correspondência. Possivelmente seu desempenho foi comprometido por algum fator, por exemplo, por não terem compreendido bem a tarefa. Contudo, durante a análise dos dados, verificamos que os resultados não eram alterados significativamente com a retirada desses sujeitos, portanto, optamos por mantê-los na análise dos dados.

Com relação aos estímulos, observou-se que alguns apresentaram taxa de identificação maior que os demais. Os que apresentaram os maiores valores foram *dali-lida* (dal), *guichê-chegue* (gui) e *cassis-cisca* (cas). Durante a análise dos dados, verificamos que a exclusão desses estímulos também não alterou significativamente os resultados, pelo que optamos por mantê-los na análise.

### 6.3.2.4 Resultados

A variável dependente neste experimento é a palavra identificada para cada estímulo: uma oxítone ou uma paroxítone – “resposta paroxítone”. A variável independente é a

frequência de ocorrência das palavras associadas ao estímulos, que foi avaliada como a “frequência relativa da palavra paroxítone”:

$$\left( \frac{\text{frequência absoluta da paroxítone}}{\text{frequência absoluta da paroxítone} + \text{frequência absoluta da oxítone}} \right)$$

Os dados obtidos foram ajustados com base em um modelo generalizado linear com efeitos mistos<sup>35</sup>, como mostrado em (i).

(i) resposta ~ frequência + (1 | sujeito)

Foi necessário adotar um modelo generalizado linear visto que a variável dependente “resposta da paroxítone” é percentual, com uma distribuição de erros binomial. A variável independente “frequência relativa da paroxítone” de cada estímulo foi considerada como um efeito fixo e contínuo no modelo. Os sujeitos foram considerados como um efeito aleatório no modelo, o que equivale a assumir que cada sujeito apresenta uma tendência particular quanto à proporção de escolha da forma paroxítone.

A seguir, a FIG. 31 mostra a distribuição dos dados obtidos e os dados ajustados ao modelo linear.

---

<sup>35</sup> R: função *lmer*, pacote *lme4*

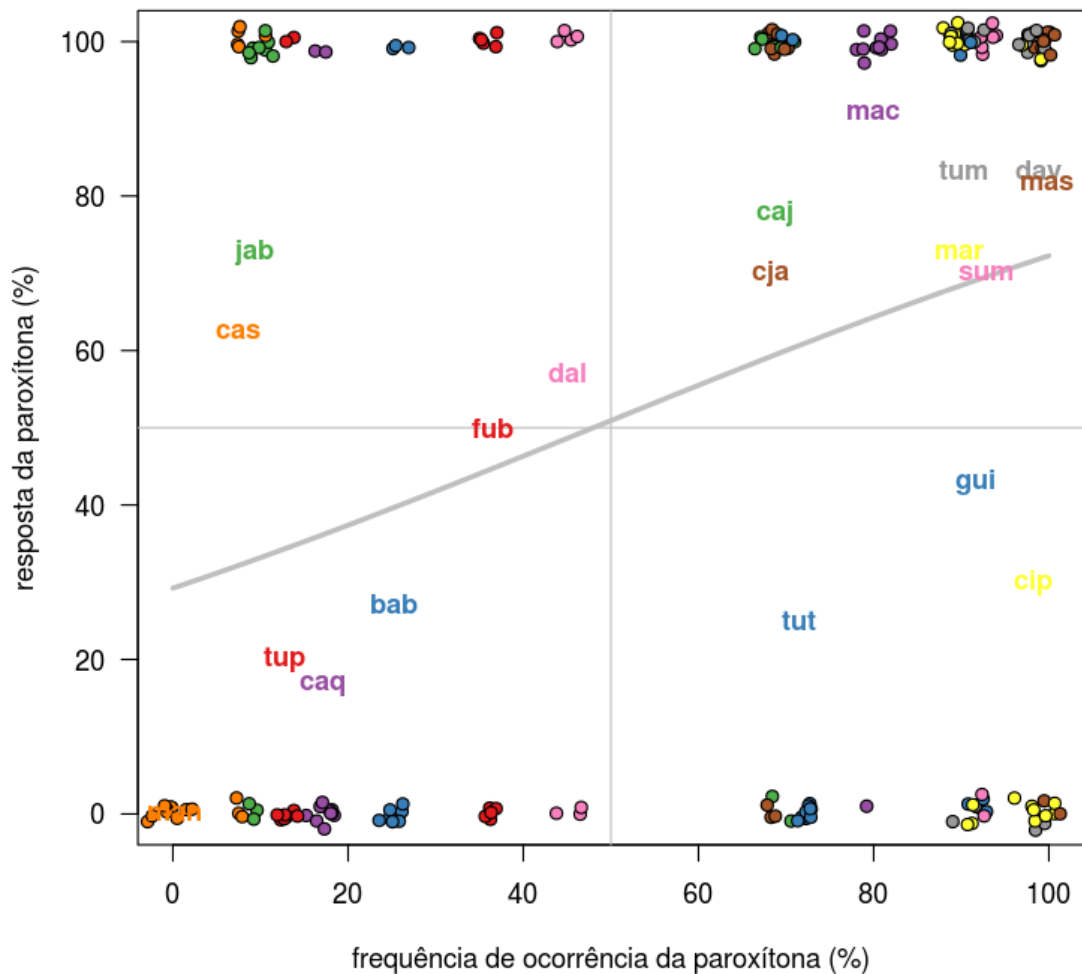


FIGURA 31 - Percentual de respostas paroxítonas em função da frequência de ocorrência relativa das paroxítonas

man (manoel-noelma)	mar (marrom-roma)	cas (cassis-cisca)
dav (davi-vida)	mac (macau-calma)	sum (sumô-moço)
caj (caju-juca)	jab (jabuti-botija)	caq (caqui-quica)
tum (tumor-morto)	fub (fubá-bafo)	mas (maçom-soma)
dal (dali-lida)	tut (tutor-torto)	cja (cajá-jaca)
tup (tupi-pito)	cip (cipó-posse)	
bab (babá-baba)	gui (guichê-chegue)	

Na FIG. 31, o eixo horizontal indica a frequência de ocorrência relativa da palavra paroxítona associada aos estímulos. O eixo vertical indica a porcentagem de respostas paroxítonas. As respostas dos sujeitos para cada estímulo correspondem aos pontos plotados em 0% (resposta = palavra oxítona) ou em 100% (resposta = palavra paroxítona) de resposta à

paroxítona. Note-se que as cores dos estímulos se repetem. Os estímulos foram representados em cores diferentes e os pontos foram ligeiramente espaçados, para melhor visibilidade. Os códigos de três letras (as três primeiras letras da palavra oxítona do estímulo) representam a resposta média de cada estímulo. A reta em cinza é a curva de regressão obtida no ajuste ao modelo.

A FIG. 31 ilustra que a curva de regressão apresenta inclinação positiva e é significativamente diferente de zero ( $z = 4.26$ ,  $p < 0.0001$ ). Contudo, ao contrário do que havia sido proposto na hipótese inicial, não pode ser identificada qualquer tendência de preferência por paroxítonas, uma vez que, quando a frequência relativa da paroxítona é igual a 50%, a resposta da paroxítona é 51%, que não é significativamente diferente de 50% ( $z = 0.24$ ,  $p > 0.81$ ). Se houvesse uma tendência a favor das paroxítonas, seria esperada uma resposta à paroxítona significativamente maior que 50% quando a frequência relativa da paroxítona fosse 50% – ou seja, seria esperado que a reta de correlação fosse deslocada do centro para a parte superior do gráfico. Na avaliação do efeito aleatório do modelo<sup>36</sup>, “tendência por informante”, foram obtidos valores nulos, o que indica que tal efeito não é significativo. Pode-se observar uma grande dispersão dos resultados ao redor da curva de regressão, com destaque para os estímulos *jabuti-botija* (jab), *cassis-cisca* (cas), *tutor-torto* (tut), *guichê-chegue* (gui) e *cipó-posse* (cip), que se encontram muito deslocados da curva.

#### 6.3.2.5 Discussão dos resultados

Neste experimento, avaliamos a percepção de sequências sonoras linguísticas ambíguas. Procuramos testar a hipótese de que, nos padrões linguísticos acentuais, a estrutura percebida sofreria influência da frequência de ocorrência e de tipo das palavras testadas. Os resultados obtidos neste experimento confirmam parcialmente tal hipótese: os resultados indicam que a percepção do estímulo ambíguo como paroxítona ou oxítona depende da frequência relativa de ocorrência das palavras paroxítona e oxítona associadas ao estímulo. Tal relação pode ser interpretada como uma evidência positiva de que o uso linguístico influencia o acesso lexical e de que, portanto, informações sobre a frequência de ocorrência das palavras devem ser armazenadas no léxico, indo ao encontro de diversos estudos já

---

<sup>36</sup> R: função *ranef*, pacote *lme4*

mencionados no cap. 3. Por outro lado, os resultados não evidenciam nenhuma influência da frequência de tipo – que equivaleria a uma preferência absoluta pela escolha de paroxítonas –, ao contrário do que previa a hipótese inicial. O experimento, por si só, não descarta, contudo, a existência de tal tendência.

Os resultados apresentados conseguem apreender um efeito estatístico que é global para a população – a correlação entre a identificação da palavra e a frequência de ocorrência. Observa-se, contudo, uma grande dispersão nos resultados, de forma que alguns estímulos se afastam mais que os outros da curva de correlação (cf. FIG. 31). Esse fato pode ser devido a uma variação maior nas frequências próprias de cada informante para tais estímulos, que diferem de algum modo da frequência indicada pelo *corpus* de referência utilizado. O efeito aleatório sujeito não ter sido significativo indica que o experimento não foi capaz de identificar a existência de tendências pessoais para os sujeitos. Não descarta, todavia, a existência de tais tendências.

Reverendo todo o paradigma experimental, foi identificada uma circunstância que pode ter contribuído para a não significância do efeito da paroxítona. Na manipulação das propriedades acústicas das sílabas para construção dos estímulos, foi utilizada como referência a proporção entre a sílaba forte e a sílaba fraca de uma palavra gravada separadamente. Esse procedimento, entretanto, ainda que tenha promovido o controle experimental, pode ter contribuído para que não fosse possível evidenciar uma preferência à escolha de paroxítonas: não foi considerada a diferença na duração e intensidade entre a sílabas fracas pretônica (da oxítone) e a postônica (da paroxítone). Como mostrado pelos resultados do experimento de produção, as sílabas pretônicas apresentam duração maior que as postônicas. Ou seja, a sílaba átona de uma oxítone é mais longa e mais intensa que a de uma paroxítone. Na manipulação das sílabas, foram usados valores de referência para duração e intensidade que, em média, se aproximam mais aos de uma pretônica. Como resultado, a sílaba fraca não foi suficientemente reduzida em duração e intensidade. Com isso, os estímulos podem ter se tornado tendenciosos à percepção de oxítonas e mascarado uma possível preferência global pela percepção das paroxítonas, que era esperava pela hipótese inicial.

A fim de solucionar o problema da construção dos estímulos e, ao mesmo tempo, possivelmente revelar uma tendência à percepção da paroxítone, seria necessário repetir o experimento, mas com novos estímulos, mais “ecológicos”.

### 6.3.3 Experimento 3: percepção categórica

Este último experimento de percepção teve como objetivo investigar a categorização de padrões acentuais. Como vimos no cap. 4, a percepção categórica é um fenômeno presente em diversas esferas da linguagem humana, dos sons aos significados, sempre que se verifica uma mudança abrupta na percepção de padrões contínuos. A percepção categórica está em jogo, por exemplo, na tarefa de reconhecer palavras, uma vez que se trata de buscar e acessar categorias lexicais. Uma das formas de se testar em que nível uma tarefa perceptual é categórica envolve testes de identificação: sujeitos são submetidos a estímulos isolados e devem classificá-los ou nomeá-los. A partir do ajuste logístico dos dados obtidos no teste, é traçada, para cada sujeito, uma curva psicométrica, que modela a probabilidade de resposta como função do parâmetro físico de interpolação dos estímulos. As principais medidas realizadas na curva ajustada são intercepto e inclinação, que refletem, respectivamente, o ponto no contínuo dos estímulos onde ocorre a transição de percepção entre as categorias e quão abrupta é a transição de categorias (cf. cap. 4 para maiores detalhes sobre essas medidas).

Retomando a hipótese inicialmente levantada neste estudo de que o acento é lexicalmente especificado, decorre que, se ele é parte do conhecimento que o falante dispõe acerca de uma palavra, é, portanto, determinante para seu reconhecimento. A relação de proeminência estabelecida pelas sílabas acentuadas – ao lado de informações segmentais, semânticas, contextuais e mesmo outras, de natureza não linguística – seria um dos fatores que guiam o acesso lexical. A hipótese da especificação no léxico pode ser, assim, expandida, prevendo que, na ausência de informações sobre a estrutura sonora, o reconhecimento das palavras poderia estar sujeito a outros fatores, como o contexto ou a frequência de ocorrência, como defendem os Modelos de Uso. O mesmo é esperado quando as informações sobre o acento (e segmentais) se tornam maximamente ambíguas (e.g. se for gerado artificialmente um estímulo acústico que mescle em igual medida uma palavra oxítona, como *cará*, e uma palavra paroxítona, como *cara*). Essa hipótese subsidiária pode ser avaliada por um teste de identificação, utilizando como estímulos graus diferentes de mesclagem entre duas palavras que diferem apenas pela localização do acento (como *cará* e *cara*). Espera-se que, se as duas palavras mescladas apresentarem valores de frequência de ocorrência semelhantes, a transição

entre as categorias lexicais na curva psicométrica tenda a se localizar ao redor do nível mais intermediário dos estímulos, correspondente à forma acusticamente mais ambígua. Ao contrário, quando os valores de ocorrência das palavras mescladas são significativamente distintos, espera-se que a curva psicométrica reflita tal diferença com um deslocamento do intercepto, havendo um maior número de estímulos correspondentes à categoria mais frequente antes da mudança abrupta. Ou seja, é como se a categoria com maior frequência de ocorrência “empurrasse” a curva psicométrica e se estendesse pelos estímulos mais ambíguos, de forma a conquistar maior espaço perceptual, sendo percebida por mais tempo que a categoria menos frequente. A hipótese específica a ser testada neste experimento é a de que, em um conjunto de estímulos progressivamente ambíguos, o espaço perceptual das palavras associadas a tal conjunto pode ser determinado a partir de sua frequência de ocorrência.

#### 6.3.3.1 Sujeitos

Foram testados 22 sujeitos, estudantes de nível universitário, voluntários, de 20 a 30 anos, igualmente divididos por sexo. Todos os sujeitos são falantes nativos do português brasileiro e não relataram qualquer história de problemas de fala ou audição.

#### 6.3.3.2 Estímulos

Os estímulos testados consistiram de 10 pares mínimos quanto ao acento lexical (i.e. duas palavras diferentes com sequências sonoras idênticas, exceto pela posição do acento, e.g. “tato”, “tatu”). Os pares mínimos selecionados foram manipulados por meio do programa de *morphing* acústico Morphala (LABOISSIÈRE et al., 2012), que tem como grande vantagem a naturalidade dos padrões sonoros gerados, ao lado do controle preciso e objetivo das variáveis manipuladas. De forma simplificada, o *morphing* opera mesclando progressivamente, por interpolação, os dois sons originais, gerando “híbridos” distribuídos de forma uniforme entre os estímulos extremos. Também os níveis extremos são manipulados, de forma a homogeneizar os estímulos quanto a possíveis imperfeições decorrentes do processo de

filtragem e ressíntese envolvidos na manipulação. Cada “híbrido”  $H$  corresponde a um nível da interpolação entre o extremo  $A$  e o extremo  $B$ , em  $n$  níveis – podendo ser descrito como

$$H = \frac{(n - i)A + iB}{n} \quad \text{onde } i = 0, 1, 2, \dots, n$$

O *Morphala* foi utilizado para interpolar as duas palavras de cada par mínimo em onze níveis, resultando em 110 estímulos: 11 níveis x 10 pares de palavras. O diagrama apresentado na FIG. 32 resume a construção e organização dos estímulos experimentais.

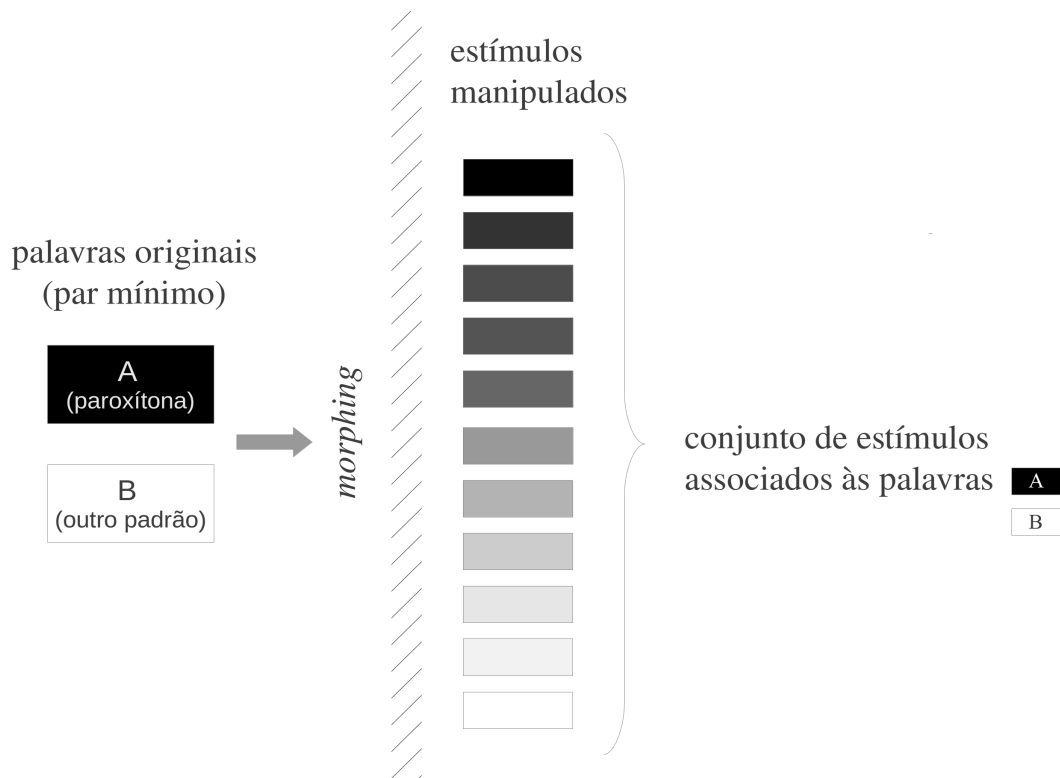


FIGURA 32 - Diagrama esquemático do processo de construção dos estímulos do experimento de percepção 3

Como se pode ver na FIG. 32, cada par mínimo selecionado é composto por duas palavras que diferem apenas quanto à posição do acento: palavra  $A$  e palavra  $B$ . Tais palavras foram submetidas ao *morphing*, de forma a gerar, em níveis progressivos de semelhança, 11 estímulos manipulados, que formam, neste experimento, o conjunto de estímulos associados às palavras  $A$  e  $B$ .

As FIG. 33a, 33b e 33c, a seguir, mostram exemplos de estímulos gerados pelo *morphing*, sendo os extremos “fábrica” e “fabrica” contrapostos aos níveis dois, quatro, seis e oito de interpolação.

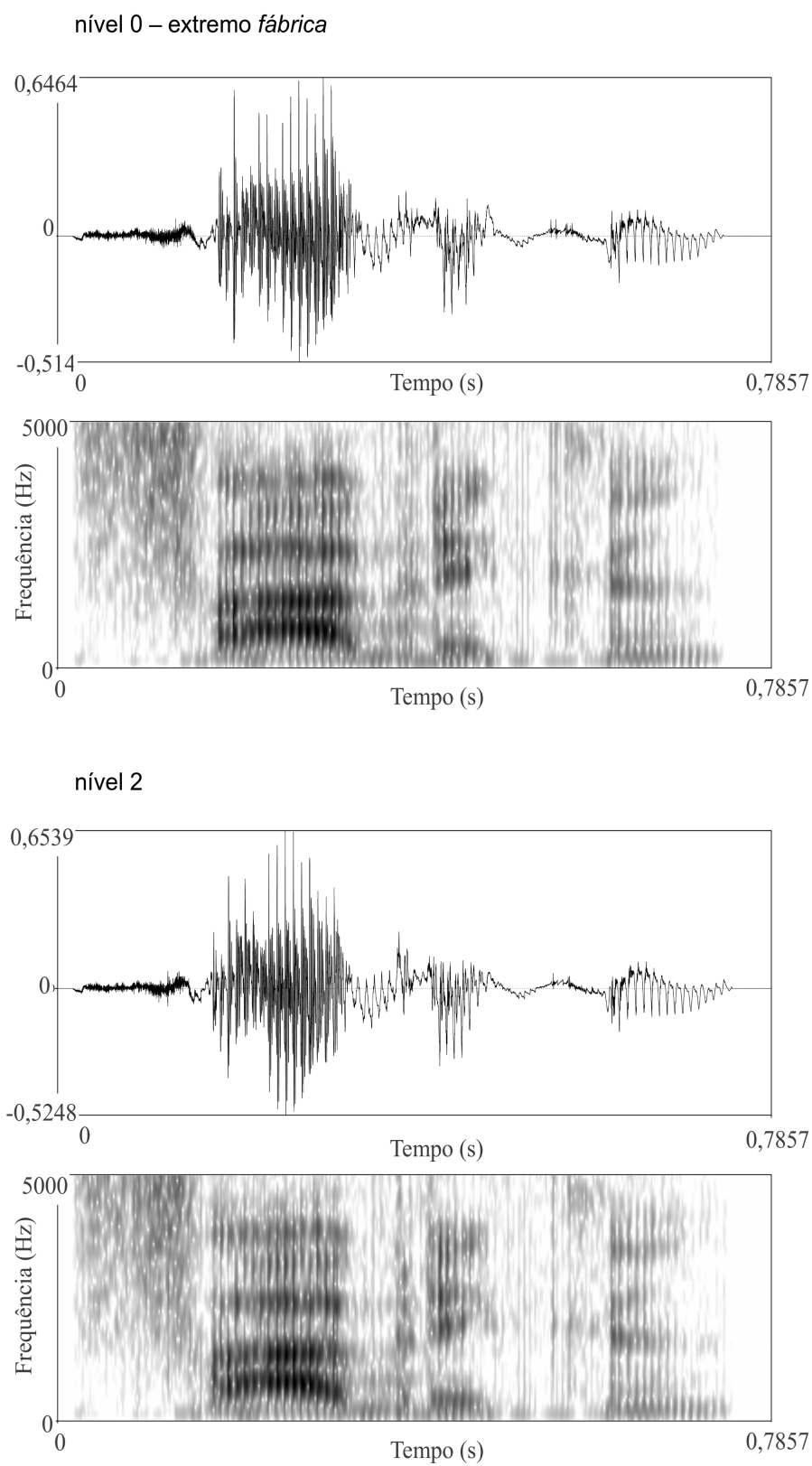


FIGURA 33a - Exemplos de estímulos gerados pelo *morphing* no experimento de percepção 3  
conjunto de estímulos *fábrica-fábrica*, níveis 0 (extremo *fábrica*) e 2

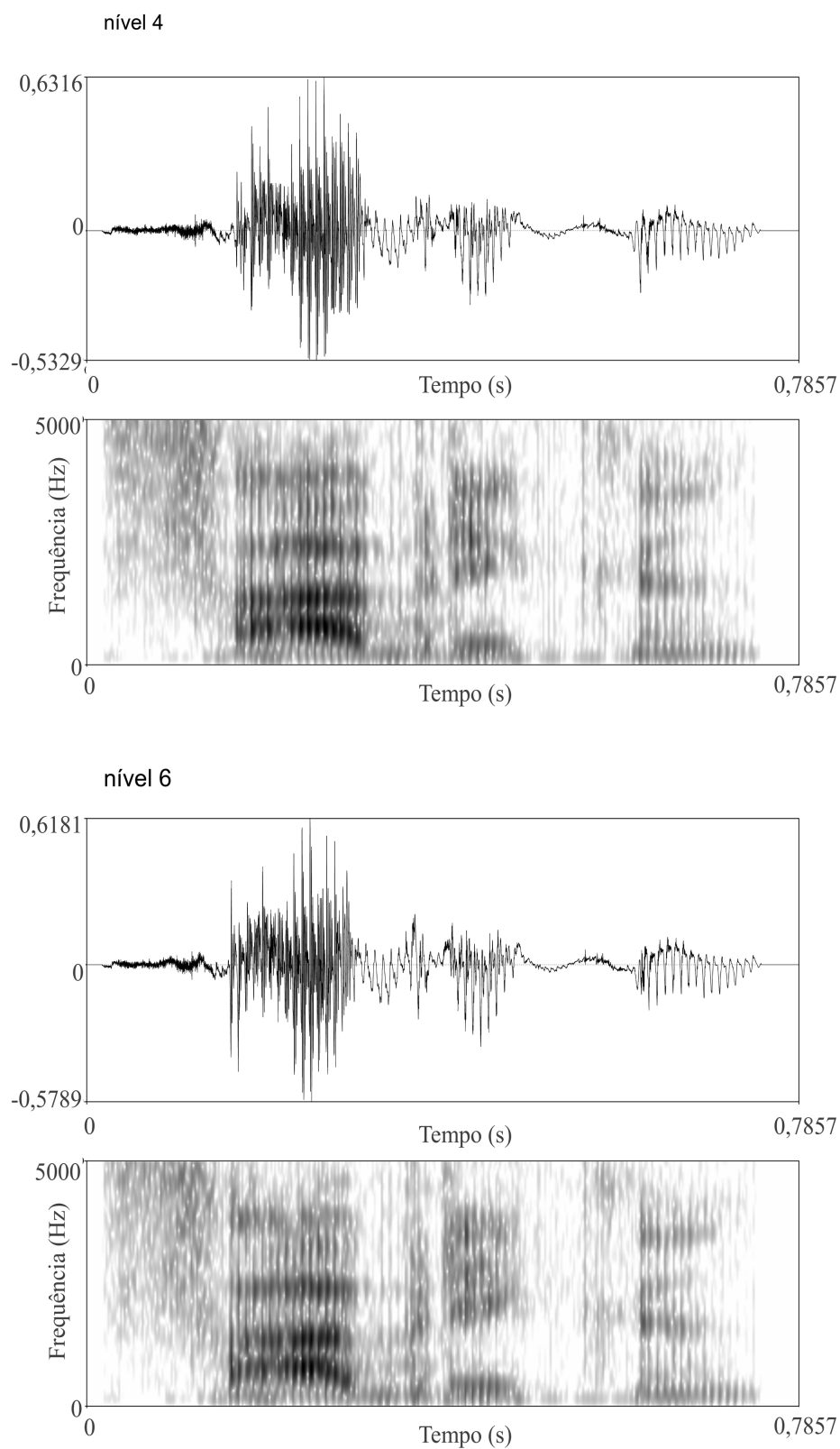


FIGURA 33b - Exemplos de estímulos gerados pelo *morphing* no experimento de percepção 3  
conjunto de estímulos *fábrica-fábrica*, níveis 4 e 6

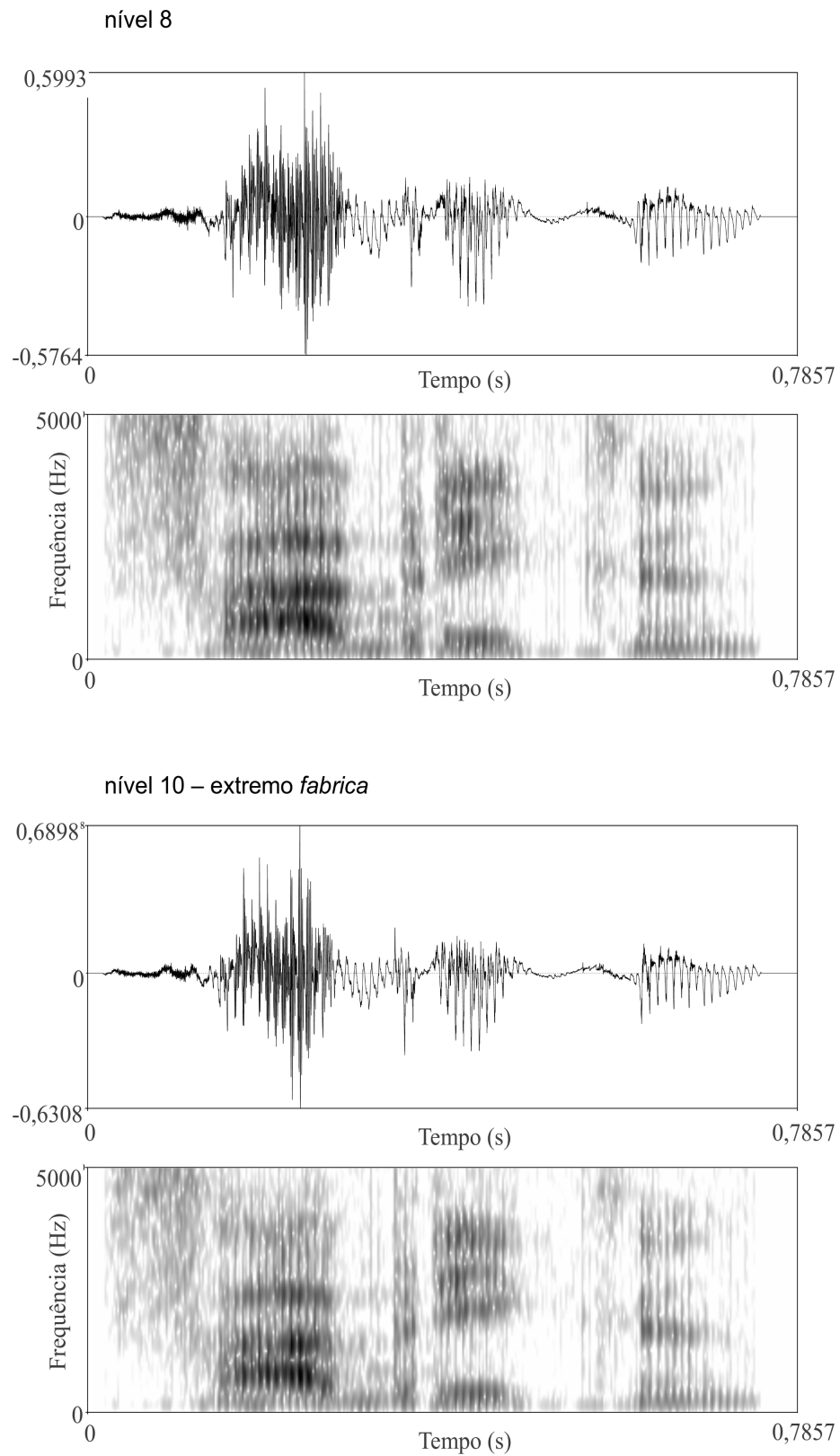


FIGURA 33c - Exemplos de estímulos gerados pelo *morphing* no experimento de percepção 3  
conjunto de estímulos *fabrica-fábrica*, níveis 8 e 10 (extremo *fabrica*)

A seguir, o QUADRO 9 apresenta os pares de palavras associadas aos conjuntos de estímulos, assim como seus respectivos valores de frequência de ocorrência, obtidos no Projeto Aspa (cf. cap. 5).

<b>Palavra 1</b>	<b>Frequência de Ocorrência</b>	<b>Palavra 2</b>	<b>Frequência de Ocorrência</b>
bate	11056	bati	278
pito	206	pitu	29
come	2879	comi	265
tato	441	tatu	388
ficaram	22944	ficarão	3738
secretária	9941	secretaria	43094
publica	3881	pública	52695
hesito	32	êxito	4106
fabrica	1652	fábrica	18455
pratica	1638	prática	27106

QUADRO 9 - Palavras associadas aos conjuntos de estímulos utilizados no experimento de percepção 3 e seus valores de ocorrência

Cada conjunto de estímulos é associado a um par de palavras, sendo uma delas sempre paroxítona e a outra podendo ser oxítona ou proparoxítona. Por essa razão, a palavra paroxítona associada ao conjunto foi adotada como referência para o conjunto de estímulos. Note-se que encontramos dificuldades em determinar qual das palavras do par *secretária-secretaria* seria a paroxítona, uma vez que a sequência de vogais no fim das duas palavras pode ser pronunciada tanto como ditongo quanto como hiato. Optamos provisoriamente por classificar *secretária* como proparoxítona e *secretaria*, como paroxítona (i.e. interpretando a sequência de vogais como um hiato em ambas as palavras).

### 6.3.3.3 Procedimentos

Os 110 estímulos sonoros originados pelo *Morphala* foram apresentados um a um, de forma aleatória, juntamente com uma tela contendo as duas palavras associadas ao estímulo,

por meio do DmDX<sup>37</sup> (FORSTER; FORSTER, 2003). A cada estímulo ouvido, o sujeito deveria selecionar uma entre duas palavras apresentadas na tela do computador, em uma tarefa de decisão forçada. A localização de cada palavra à esquerda ou à direita da tela também foi determinada aleatoriamente. A seleção das palavras pelos sujeitos era feita apertando-se a tecla “e” ou a tecla “p” do teclado – sobre as quais foi afixado um quadrado colorido –, sendo “e” correspondente à palavra à esquerda e “p” à palavra da direita. A apresentação dos estímulos foi precedida por um breve treinamento, que consistiu em apresentar a gravação original de palavras de cinco pares mínimos diferentes dos testados, seguindo o mesmo paradigma de apresentação, mas indicando ao sujeito qual tecla apertar. O treinamento tinha como objetivo habituar o sujeito à tarefa do experimento: escutar um som, identificar a palavra ouvida entre as duas palavras apresentadas na tela do computador e apertar o botão do teclado correspondente à palavra identificada. Na metade dos estímulos em teste, era realizada uma pausa regulada pelos próprios sujeitos e que variou de um a dois minutos e meio. Neste experimento, os sujeitos levaram cerca de 30 minutos para completar a tarefa solicitada.

#### 6.3.3.4 Resultados

A variável dependente no experimento são as respostas como oxítone ou paroxítone. Foi inserida no modelo como a resposta à palavra com maior valor de frequência de ocorrência. Por exemplo, no conjunto de estímulos *pratica-prática*, as respostas foram computadas em relação a *prática*, que é mais frequente que *pratica*. As variáveis independentes são (1) o nível de *morphing* do estímulo, avaliado no modelo como valores de 0 a 1 referentes aos 11 níveis de *morphing* e (2) a frequência de ocorrência das palavras associadas ao estímulos, avaliada no modelo como a frequência relativa da palavra paroxítone (calculada como a frequência absoluta da paroxítone dividida pela frequência absoluta da outra palavra do par, fosse ela oxítone ou proparoxítone).

---

<sup>37</sup> Programa desenvolvido por Jonathan Forster para apresentação de estímulos auditivos e visuais e para medição de tempo de reação em experimentos de psicologia. Disponível em <http://www.u.arizona.edu/~kforster/dmdx/dmdx.htm>.

Os dados obtidos foram ajustados a partir de um modelo linear generalizado com efeitos mistos<sup>38</sup>, tendo dois fatores fixos – intercepto e inclinação globais – e dois fatores aleatórios – sujeito e conjunto de estímulos –, como representado em (i).

(i) respostas  $\sim$  morphing + (1 + morphing | sujeito) + (1 + morphing | estímulo)

Tal modelo assume que cada sujeito e cada conjunto de estímulos apresentam intercepto e inclinação próprios. As quatro possibilidades de combinação entre fatores fixos e aleatórios foram verificadas individualmente – comparando-se os modelos reduzidos, contendo as diferentes combinações de fatores, e o modelo completo, contendo os quatro fatores –, por meio do teste de razão de verossimilhança<sup>39</sup>, que respeita a distribuição de qui-quadrado:

(inclinação | sujeito):  $\chi^2[2] = 35,4$   $p < 0,0001$

(inclinação | estímulo):  $\chi^2[2] = 1,7$   $p = 0,42$

(intercepto | sujeito):  $\chi^2[2] = 45,3$   $p < 0,0001$

(intercepto | estímulo):  $\chi^2[2] = 30,0$   $p < 0,0001$

Como se pode ver, todas as combinações de fatores se mostraram significativas a pelo menos  $p < 0,0001$ , exceto a que toma a combinação (inclinação | estímulo), que foi retirada do modelo completo, resultando no seguinte modelo final:

(ii) respostas  $\sim$  morphing + (1 + morphing | sujeito) + (1 | estímulo)

Ao analisar os dados, verificou-se uma resposta anômala a dois pares de estímulos, *bate-bati* e *secretaria-secretaria*, que apresentaram valores muito discrepantes nos extremos, em que se esperam valores próximos de 0% e 100%. As FIG. 34 e 35, na sequência, mostram os valores de respostas de cada par de estímulos nos extremos, respectivamente, 0% e 100%.

---

38 R: função *glmer*, pacote *lme4*

39 R: função *anova*, pacote *stats*

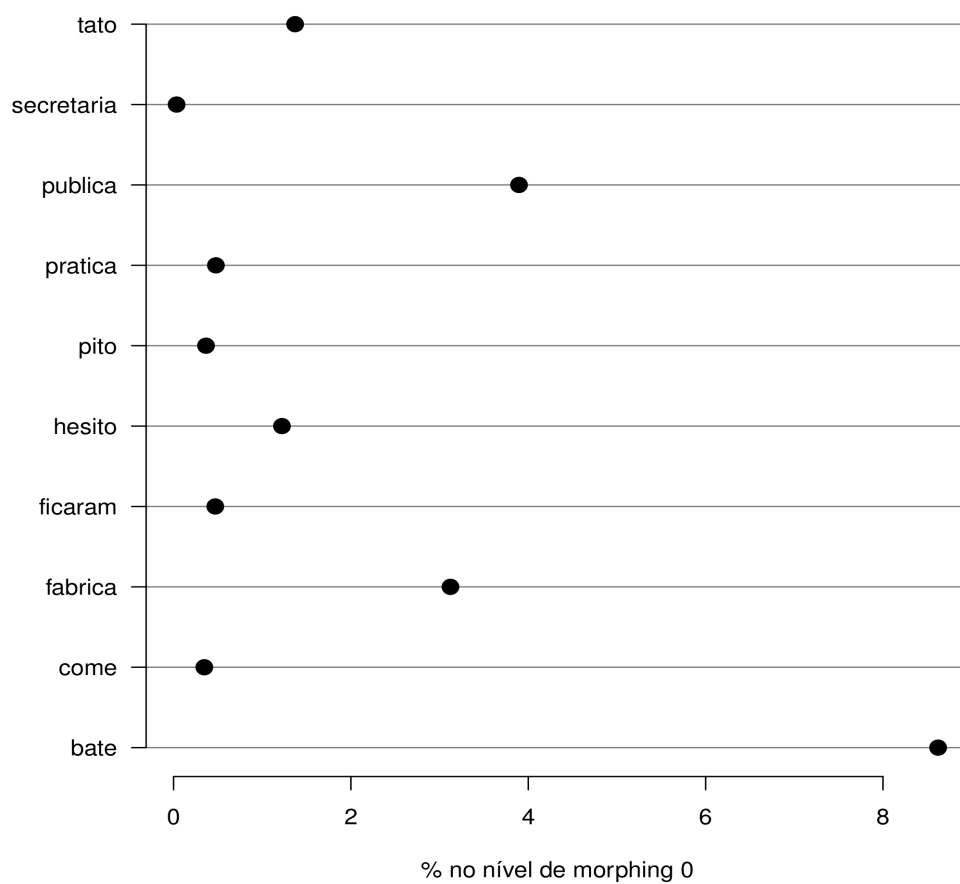


FIGURA 34 - Valores das respostas a cada par de estímulos no extremo 0%

A FIG. 34 mostra que os pares de palavras testados apresentam valores distintos de resposta no nível extremo de *morphing* a 0%, em que seria esperada uma aglomeração das respostas em torno do 0.

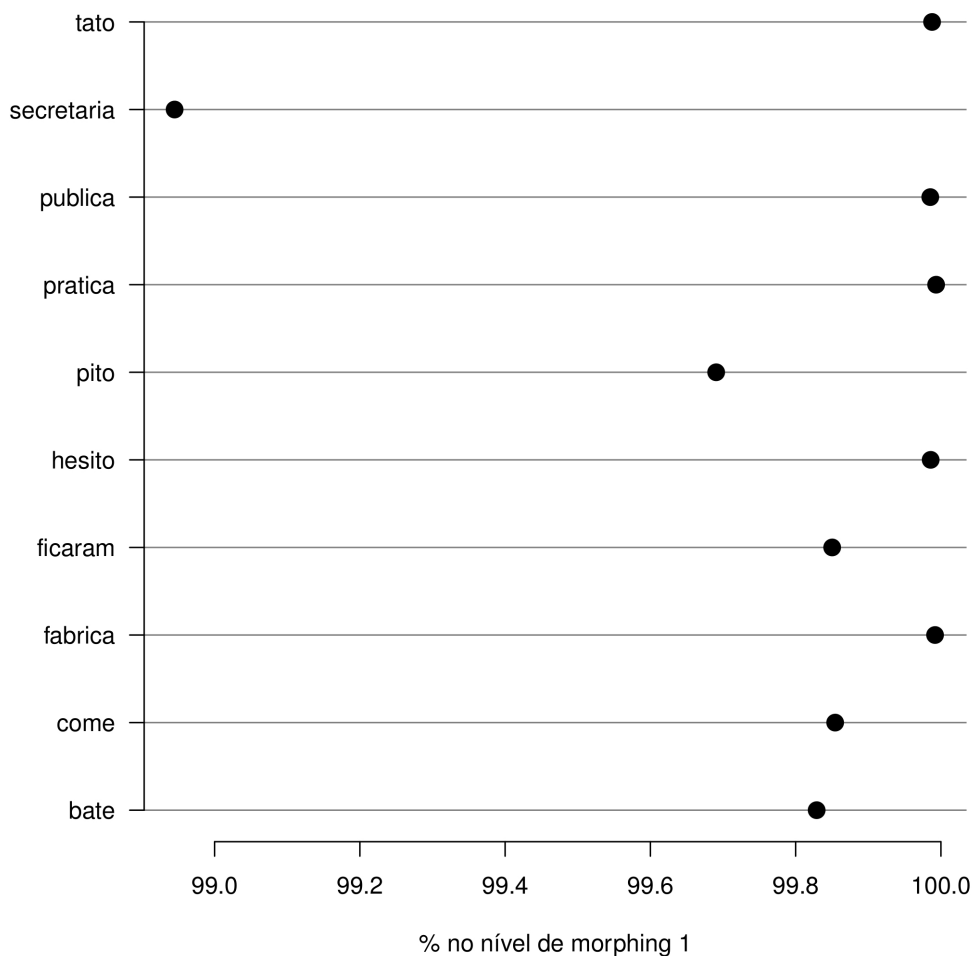


FIGURA 35 - Valores das respostas a cada par de estímulos no extremo 100%

A FIG. 35 mostra que os pares de palavras testados apresentam valores distintos de resposta no nível extremo de morphing a 100%, em que seria esperada uma aglomeração das respostas em torno do 100.

Diante da variabilidade nas respostas de diferentes estímulos junto aos níveis extremos de *morphing*, decidiu-se por eliminar quaisquer pares de estímulos que não atingissem 5% ou 95% nos extremos. Apenas o par *bate-bati* não passou no critério e foi eliminado. O modelo foi, então, reajustado sem esse par.

A frequência de ocorrência não foi inserida como fator fixo para evitar a inflação dos graus de liberdade dos resíduos. O tratamento dessa variável independente será realizado na etapa imediatamente seguinte.

Foram então obtidos os coeficientes dos fatores fixos e aleatório<sup>40</sup>. A partir do coeficiente do fator aleatório, o intercepto estimado para cada conjunto de estímulos foi confrontado com a frequência de ocorrência relativa da palavra paroxítona associada ao conjunto de estímulos correspondente, por meio de um teste Pearson de correlação<sup>41</sup>. Foi obtido um coeficiente de correlação  $R = 0,78$ , que é significativamente diferente de zero ( $t[7] = 3,28$ ,  $p < 0,05$ ). Pode-se concluir, portanto, que há uma correlação entre as duas variáveis: intercepto específico a cada conjunto de estímulos e frequência de ocorrência relativa da paroxítona associada, como se pode observar na FIG. 36 abaixo.

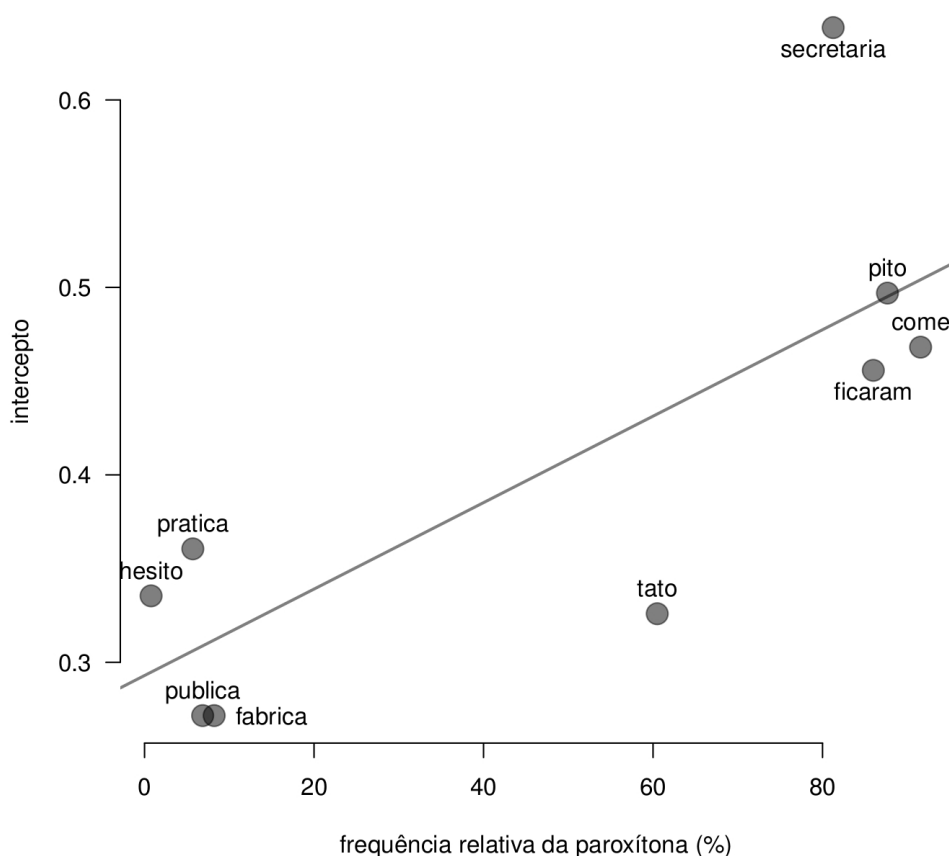


FIGURA 36 - Correlação entre intercepto dos conjuntos de estímulos e frequência de ocorrência relativa da paroxítona ( $R = 0,78$ ,  $t[7] = 3,28$ ,  $p < 0,05$ )

Como ilustrado na FIG. 36, há uma correlação entre a localização do intercepto das respostas para cada par de palavras – i.e., a transição entre categorias – e a frequência de ocorrência relativa dos itens.

40 R: funções *fixef* e *ranef*, respectivamente, ambos no pacote *lme4*

41 R: função *cor.test*, pacote *stats*

No Apêndice C, as FIG. 38 a 46 mostram, para cada conjunto de estímulos, as curvas psicométricas correspondentes às respostas individuais de cada um dos sujeitos. Por fim, a FIG. 37, a seguir, ilustra o fator aleatório sujeito.

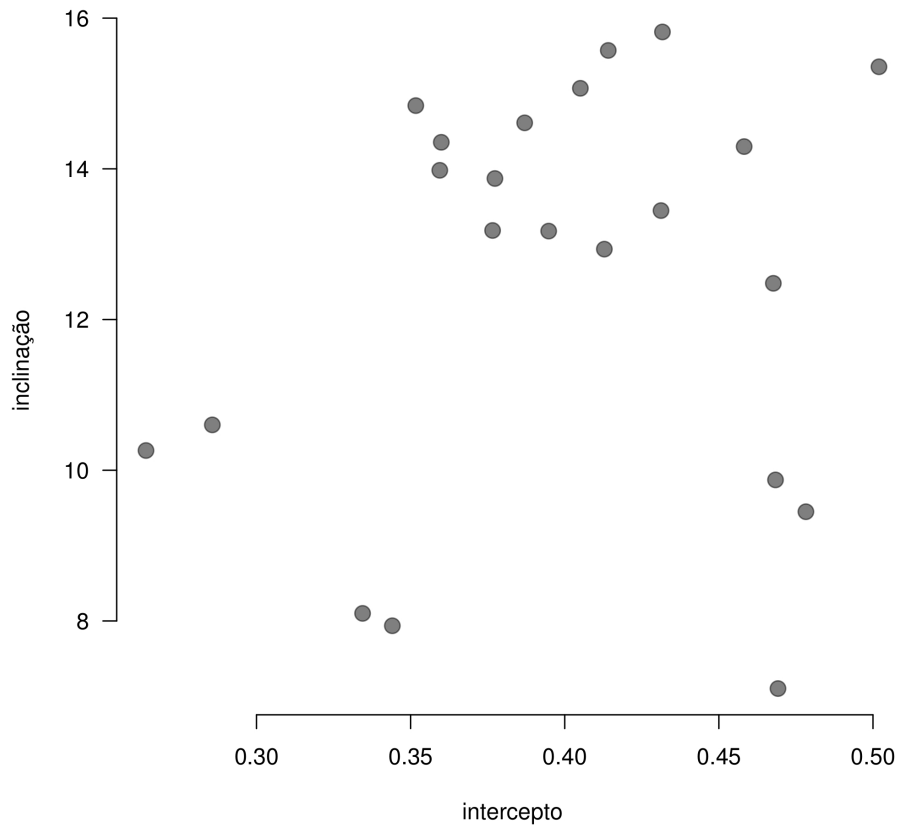


FIGURA 37 - Inclinação vs. intercepto do fator aleatório sujeito

A FIG. 37 mostra a distribuição dos sujeitos considerando a inclinação e o intercepto de suas respostas. Não pode ser observada nenhuma tendência, o que corrobora a adoção dos sujeitos como efeito aleatório no modelo.

#### 6.3.3.5 Discussão dos resultados

Este experimento avaliou a percepção categórica de padrões acentuais no PB. A hipótese inicial deste terceiro experimento previa uma influência da frequência de ocorrência

na identificação de palavras em um conjunto de estímulos progressivamente ambíguos. Os resultados levantados experimentalmente dão suporte a essa hipótese, uma vez que se verificou uma correlação entre o intercepto particular a cada conjunto de estímulos e a frequência de ocorrência das palavras associadas ao conjunto de estímulos. Tal resultado tem duas grandes consequências em termos de implicações teóricas. Em primeiro lugar, pode ser interpretado como evidência empírica da influência do uso linguístico na percepção da fala, como atestado na literatura. Em segundo lugar, serve também como evidência de que o acento deve ser lexicalmente especificado, uma vez que, na ausência de informações sobre a estrutura sonora, uma informação lexicalmente armazenada – a frequência de uso – foi capaz de influenciar a identificação das palavras.

Um resultado surpreendente foram as tendências de favorecimento de uma das palavras do par de estímulos. A nos basearmos em um possível efeito da frequência de tipo, seria esperado que as palavras pertencentes ao tipo de maior frequência – as paroxítonas – fossem privilegiadas frente à outra palavra do par de estímulos e fossem percebidas antes, contando, assim, com maior espaço perceptual. Entretanto, os resultados indicam que as paroxítonas parecem ser desfavorecidas com relação às oxítonas e proparoxítonas. Retomando a FIG. 31, vemos que os pares de palavras em que a paroxítona é menos frequente apresentam menor intercepto que os pares em que a paroxítona é mais frequente. É claro que, devido ao número reduzido de palavras, não é possível afirmar se tal comportamento é geral. A interpretação de tais resultados não é simples e talvez não seja possível explicar exatamente os mecanismos por trás do desfavorecimento das paroxítonas.

A tentativa de explicação que proporemos recorre a características fonéticas das sílabas acentuadas e átonas, bem como a processos fonológicos observados na variedade de Minas Gerais, de onde são os sujeitos do experimento e o locutor dos estímulos.

Os pares de palavra em que se observou um desfavorecimento das paroxítonas são formados por paroxítona-proparoxítona (e.g. *pratica-prática*). Como sabemos do experimento de produção, as sílabas pretônicas são acusticamente mais semelhantes às acentuadas (em termos de valores de duração, intensidade e F0) do que as postônicas. Em uma palavra proparoxítona, temos duas sílabas postônicas, o que, por extensão, tornaria paroxítonas acusticamente mais semelhantes às oxítonas que às proparoxítonas. Nas proparoxítonas, adicionalmente, há dois *loci* de contraste: além do contraste da acentuada com a postônica, as duas postônicas contrastam entre si. Tais características acústicas fariam a proparoxítona mais

saliente em comparação com a paroxítona associada ao mesmo par, e poderiam estar por trás do maior espaço perceptual conferido à proparoxítona, o que explicaria parcialmente nossos resultados. A outra parte da explicação servirá para justificar a preferência pelas oxítonas, em decorrência de fenômenos sonoros de mudança e variação.

Como discutimos no cap. 4, a fonologia do PB se estrutura ao redor das sílabas acentuadas. Na fala corrente, muitas vezes, devido aos inúmeros cancelamentos de vogais átonas e emergência de encontros consonantais, temos um grande índice de sílabas acentuadas na produção da fala (cf. os estudos de Souza (2012) e Leite (2006) sobre o cancelamento de vogais pretônicas e postônicas na variedade de Belo Horizonte). Em decorrência do fenômeno de cancelamento de vogais átonas finais, que tem ampla ocorrência na variedade mineira, ocorre uma reestruturação dos padrões silábicos: paroxítonas se tornam oxítonas pelo cancelamento de vogais átonas (e.g. *passé* ['pas]); proparoxítonas se tornam paroxítonas ou mesmo oxítonas, seja pelo mesmo fenômeno, seja pela formação de encontros consonantais (e.g. *príncipe* ['prĩspi]); novas proparoxítonas são formadas por epêntese (e.g. *ritmo* ['hitʃimʊ]).

Esse percurso de reestruturação pode ser interpretado, em uma perspectiva dinâmica, como decorrente da auto-organização do sistema sonoro, em que se mantêm os padrões acentuais, mas mudam as palavras associadas a cada padrão, como em uma “dança das cadeiras”. Mudanças em cadeia são recorrentes nas línguas do mundo e podem ser interpretadas como resultado da coordenação de padrões gramaticais. Uma das consequências dessa cadeia de eventos para a estrutura sonora do PB é uma super-inflação do novo padrão oxítono, que passa a abarcar também as antigas paroxítonas, mais numerosas em tipo.

A modelagem multirrepresentacional permite explicar como, diante de fenômenos de mudança e variação linguística, somos capazes de lidar simultaneamente com formas concorrentes de uma mesma palavra. Se no contínuo da fala os contrastes entre as três classes acentuais se reduzem drasticamente, contudo, as formas isoladas, associadas a uma das três classes, não se perdem, pois continuam lexicalmente armazenadas nas representações dos exemplares. A redundância das informações nos permite acessar, dessa forma, tanto a forma reduzida da palavra na fala contínua quanto a forma completa, em isolamento.

## 7 Conclusões

Este estudo teve como propósito central examinar o acento lexical no português brasileiro, a partir de uma perspectiva metodológica que conjugasse teorias fonológicas com evidências experimentais. Optou-se por uma abordagem abrangente, que considerasse tanto aspectos da produção quanto da percepção do acento.

Foram realizados um experimento de produção e três experimentos de percepção de fala. No experimento de produção, avaliamos propriedades acústicas e articulatórias envolvidas na realização do acento. Testamos cinco sujeitos de Minas Gerais, por meio da gravação simultânea do sinal da fala, do sinal eletroglotográfico e da movimentação facial. Verificamos que a principal propriedade acústica empregada na realização do acento é a duração, capaz de diferenciar sílabas acentuadas de pretônicas e postônicas e as últimas entre si. Também se mostraram relevantes F0 e intensidade, mas com menor poder de contraste. Da interação entre acento e foco, pudemos verificar que as propriedades acústicas envolvidas na produção do acento estabelecem uma relação complexa em termos da sinalização de proeminência no nível da palavra (acentos) e em nível superior a ela (foco). Considerando a articulação no nível glótico, sugerimos que estratégias individuais de implementação articulatória para alteração de F0 e intensidade podem estar em jogo. Os resultados obtidos corroboram estudos anteriores de outras variedades de português brasileiro.

Os experimentos de percepção foram elaborados com o objetivo de avaliar a interação entre proeminência acentual e acesso lexical. O experimento 1 consistiu em tarefa de desambiguação de sequências não linguísticas com base em relações de proeminência. Os resultados obtidos neste experimento indicam que as propriedades acústicas exercem influência na percepção da proeminência em padrões sonoros, em termos de tendências gerais. Sugerimos que a interação entre a força exercida pela estrutura linguística e a força pela interpretação não linguística tem como consequência a emergência de padrões complexos de resposta.

O experimento 2 envolveu tarefa de desambiguação, como o experimento 1, mas foram utilizados estímulos linguísticos – sequências de sílabas ambíguas em termos de identificação lexical. Os resultados obtidos neste experimento indicam que a percepção do estímulo ambíguo como paroxítona ou oxítona depende da frequência relativa de ocorrência

das palavras paroxítona e oxítona associadas ao estímulo. Tal relação pode ser interpretada como uma evidência positiva de que o uso linguístico influencia o acesso lexical e de que, portanto, informações sobre a frequência de ocorrência das palavras devem ser armazenadas no léxico – e com elas, o próprio acento.

Já o experimento 3 abordou a percepção categórica, testando a identificação de pares mínimos em um contínuo de sons intermediários manipulados. Os resultados levantados experimentalmente comprovam a influência da frequência de ocorrência na identificação de palavras em um contínuo de ambiguidade. Tal resultado tem duas grandes consequências em termos de implicações teóricas. Em primeiro lugar, pode ser interpretado como evidência empírica da influência do uso linguístico na percepção da fala, como atestado na literatura. Em segundo lugar, serve também como evidência de que o acento deve ser lexicalmente especificado, uma vez que, na ausência de informações sobre a estrutura sonora, uma informação lexicalmente armazenada – a frequência de uso – foi capaz de influenciar a identificação das palavras.

No cap. 5, apresentamos uma proposta de abordagem multirrepresentacional do acento no português que partia do sistema latino. Nessa proposta, a acentuação do português teria emergido com a auto-organização de variáveis do sistema acentual latino em decorrência de uma série de mudanças fonológicas desencadeadas pela perda da quantidade vocálica. Defendemos que a presença de mais de uma pista acústica para a sílaba acentuada (relação de troca) possibilitou que duração passasse a exercer um papel relevante no contraste entre formas reduzidas e não reduzidas. Em decorrência do fenômeno de cancelamento de vogais átonas finais, que tem ampla ocorrência na variedade mineira, ocorre uma reestruturação dos padrões silábicos: paroxítonas se tornam oxítonas pelo cancelamento de vogais átonas (e.g. *passé* ['pas]); proparoxítonas se tornam paroxítonas ou mesmo oxítonas, seja pelo mesmo fenômeno, seja pela formação de encontros consonantais (e.g. *príncipe* ['pɾĩspɪ]); novas proparoxítonas são formadas por epêntese (e.g. *ritmo* ['hitʃimʊ]). Esse percurso de reestruturação pode ser interpretado, em uma perspectiva dinâmica, como decorrente da auto-organização do sistema sonoro, em que se mantêm os padrões acentuais, mas mudam as palavras associadas a cada padrão. Nas palavras reduzidas, a duração opera como detalhe fonético fino no fenômeno de redução vocálica, mantendo a distinção entre as palavras: *paz* ['pas] apresenta vogal com menor duração que a vogal de *passé* ['pa:s] reduzido. A alternância deixa de ser entre presença e ausência de vogal final ou entre oxítona e paroxítona, passando a

ser determinada pela duração – de certa forma retomando a propriedade de quantidade vocálica existente anteriormente no latim.

A proposta apresentada neste estudo, de que o acento é lexicalmente especificado, adicionalmente recebe suporte de evidências de novas palavras formadas com padrões sonoros inovadores, que não podem ser explicadas recorrendo-se ao peso silábico, como *Wilson*, *modem*, *mórmon*, *pôster*, *contêiner*. Argumentamos que informações sobre a proeminência acentual no nível da palavra estão presentes no léxico, o que explicaria com sucesso os efeitos de frequência encontrados experimentalmente na percepção do acento e inovações sonoras recentes. Propusemos uma modelagem do acento como resultado de generalizações a partir de exemplares, generalizações que seriam responsáveis pelas tendências gerais na distribuição do acento no léxico e sua interação com a gramática.

Várias questões interessantes surgiram ao longo do desenvolvimento deste trabalho e não puderam ser abordadas por merecerem um tratamento extenso. Por exemplo, em análises linguísticas multirrepresentacionais, como a que aqui realizamos, torna-se necessário lidar com a complexidade das relações lexicais e sua influência na frequência de uso. Sabemos que as palavras sofrem influência dos campos lexicais a que pertencem. Em termos quantitativos, tal fator é geralmente considerado na forma de densidade da vizinhança lexical – fonológica, semântica, morfológica. Como determinar o peso a ser conferido a cada tipo de densidade? Como conjugar os efeitos exercidos pela frequência lexical de tipo e de ocorrência e pela densidade da vizinhança, em seus vários tipos?

Restam ainda a analisar os dados de movimentação facial coletados no experimento de produção de fala por meio do Optotrak. Esses resultados poderão contribuir para a compreensão dos mecanismos de coordenação de gestos e sua relação com as propriedades acústicas envolvidas na proeminência acentual.

Como palavra final, este trabalho teve o mérito de conjugar modelos teóricos diferentes dos tradicionalmente utilizados na análise do acento no PB, propondo uma abordagem original, pela modelagem em redes com princípios dinâmicos de funcionamento.

## Apêndice A

São listadas abaixo todas as sentenças utilizadas no experimento de produção.

O avô gagá tagarela está dizendo mentiras.  
Aquela gata malhada está comendo salame.  
A vila gátara nova está pedindo socorro.  
Essa maga taróloga está ganhando dinheiro.  
A málaga taça dela está soltando pedaços.  
Belas ágatas árabes estão fazendo sucesso.  
Dinorá Gatári Macedo está lançando um romance.  
Tafetá pagável à vista está trazendo clientes.  
Um vatapá garantido está guardado pra ele.  
A conta paga com multa está mostrando o atraso.  
“Agora paga-se menos” está escrito na conta.  
Nosso mapa galático está mostrando o caminho.  
O sátrapa gago cego está querendo sossego.  
A galápaga música está tocando no rádio.  
O sushi pedido pra hoje está cheirando a passado.  
Um zebu cubute malhado está correndo no pasto.  
Vovó Pipi divertida está dizendo besteiras.  
Zazá Zipide Faraco está comprando o terreno.  
Valor lipídico alto está causando problemas.  
“Mississipi dilúvico” está passando em São Paulo.  
O príncipe dito belo está fazendo maldades.  
O Eurípides ávido está comendo lasanha.  
A mesma gataria voraz está rondando o açougue.  
O mega pagamento mensal está salvando as finanças.  
Pirarucu buliçoso está faltando no rio.  
Um outro cubo rosado está servindo de dado.  
O povo cúbule bravo está sofrendo massacres.  
Nosso cuco burríssimo está errando os minutos.  
O búbuco bule verde está vazando a bebida.  
Esse súcubo mágico está fazendo intrigas.  
Denise Pidinale Sodré está vendendo o terreno.  
Um velho cubulento e senil está pedindo moedas.

## Apêndice B

Nas seções dedicadas à análise estatística dos experimentos, optamos por adotar a notação específica utilizada no R para construção dos modelos. Neste Apêndice, tal notação será brevemente explicada. Para maiores detalhes sobre implementação de fórmulas em R, remetemos o leitor a Crawley (2007), Baayen (2008) e Bates (2010).

As fórmulas em R são construída basicamente com operadores simbólicos e nomes de variáveis e fatores. Os operadores simbólicos que usamos nos modelos lineares testados neste estudo foram: “~”, “+”, “\*”, “|”.

O operador “~” equivale à expressão “em função de” ou “explicada por”. Separa uma fórmula em duas partes: à esquerda é especificada a variável de resposta (ou variável dependente); à direita, é especificada uma função linear que explica a variável resposta a partir dos termos do modelo (ou variáveis independentes). Em R, uma fórmula como  $y \sim x$ , em que  $y$  é a variável dependente e  $x$  é uma variável independente, equivale à função matemática  $y = x$ .

O operador “+” estabelece uma relação linear entre os termos do modelo. Por exemplo, uma função matemática do tipo  $y = ax_1 + bx_2 + c$ , em que  $y$  é a variável dependente,  $x_1$  e  $x_2$  são variáveis independentes,  $a$ ,  $b$  e  $c$  são constantes, pode ser representada pela fórmula  $y \sim x_1 + x_2$ .

O operador “\*” representa o cruzamento de variáveis, que equivale aos efeitos principais das variáveis mais a interação entre elas. Dessa forma, a notação  $a * b$  é o mesmo que a notação  $a + b + a:b$ , em que “:” representa a simples interação entre  $a$  e  $b$ .

O operador “|” é utilizado para representar o agrupamento de variáveis ou fatores. É usado ainda especificamente no pacote *lme4*, que permite o ajuste de modelos com efeitos mistos (ou seja, modelos contendo não apenas efeitos fixos, como os apresentados até aqui, mas também efeitos aleatórios). Nesse contexto, o operador “|” é usado para especificar termos com efeitos aleatórios. Por exemplo, a estrutura “(1 | fator)” indica que “fator” é um termo com efeito aleatório no modelo e, uma vez que “|” é precedido por “1”, determina o ajuste do intercepto com relação a “fator”. Em uma fórmula como  $y \sim x + (1 | w)$ , em que  $y$  é a variável dependente,  $x$  é uma variável independente incluída no modelo como um termo com efeito fixo e  $w$ , como um termo com efeito aleatório, assume-se que cada nível do fator “ $w$ ” apresenta uma tendência particular quanto à variável resposta “ $y$ ”.

### Apêndice C

As FIG. 38 a 46, a seguir, se referem aos dados obtidos no experimento 3 de percepção e mostram, para cada conjunto de estímulos, as curvas psicométricas correspondentes às respostas individuais de cada um dos sujeitos. O eixo horizontal corresponde aos níveis de *morphing* dos estímulos, enquanto o eixo vertical contém as duas palavras associadas ao estímulo. A curva em vermelho corresponde à resposta média dos sujeitos para o conjunto de estímulos.

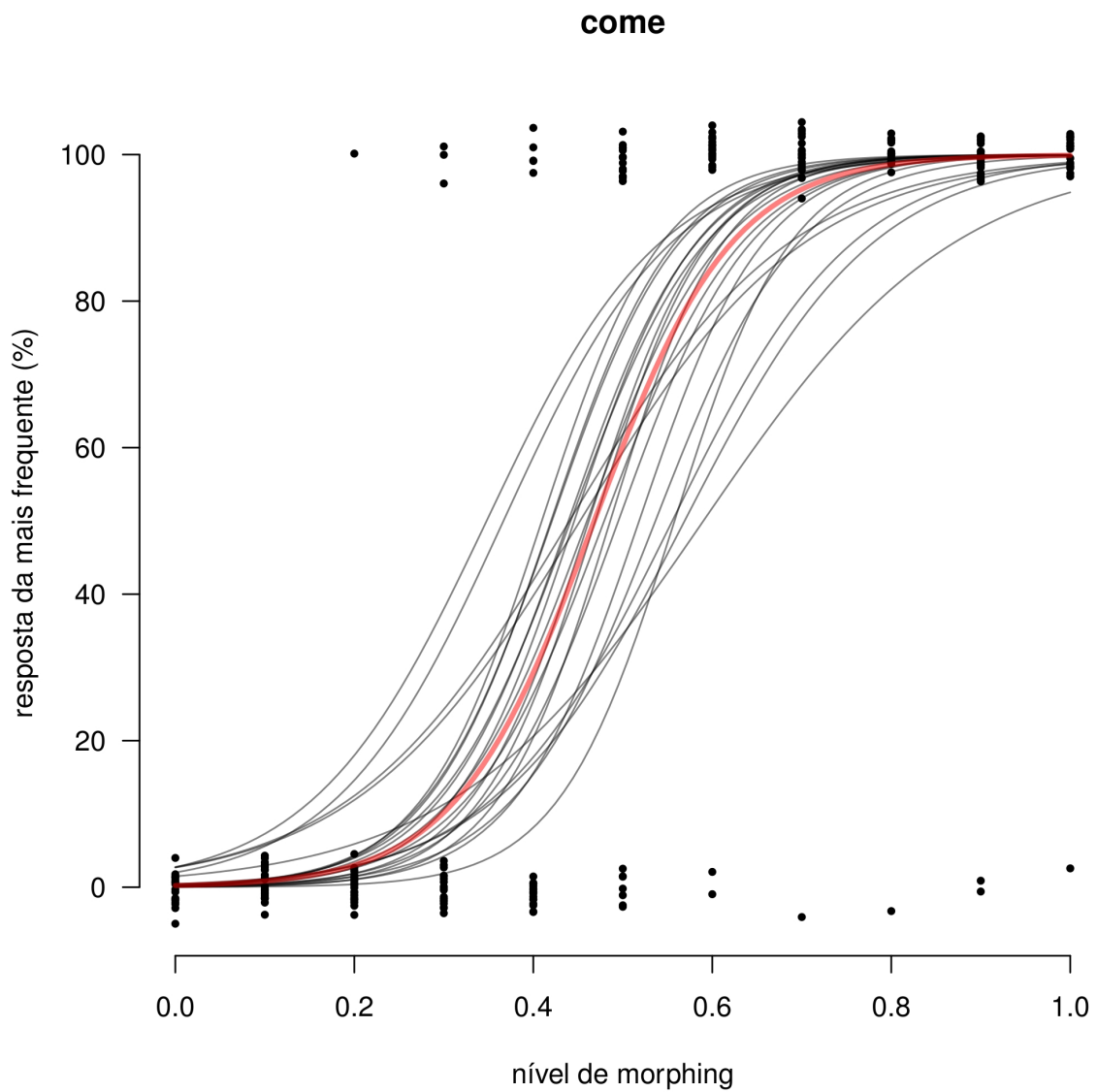


FIGURA 38 - Curva psicométrica do conjunto de estímulos *come-comi*

Eixo vertical: 0% = *comi*; 100% = *come*

Eixo horizontal: 0.0 = *comi*; 1.0 = *come*

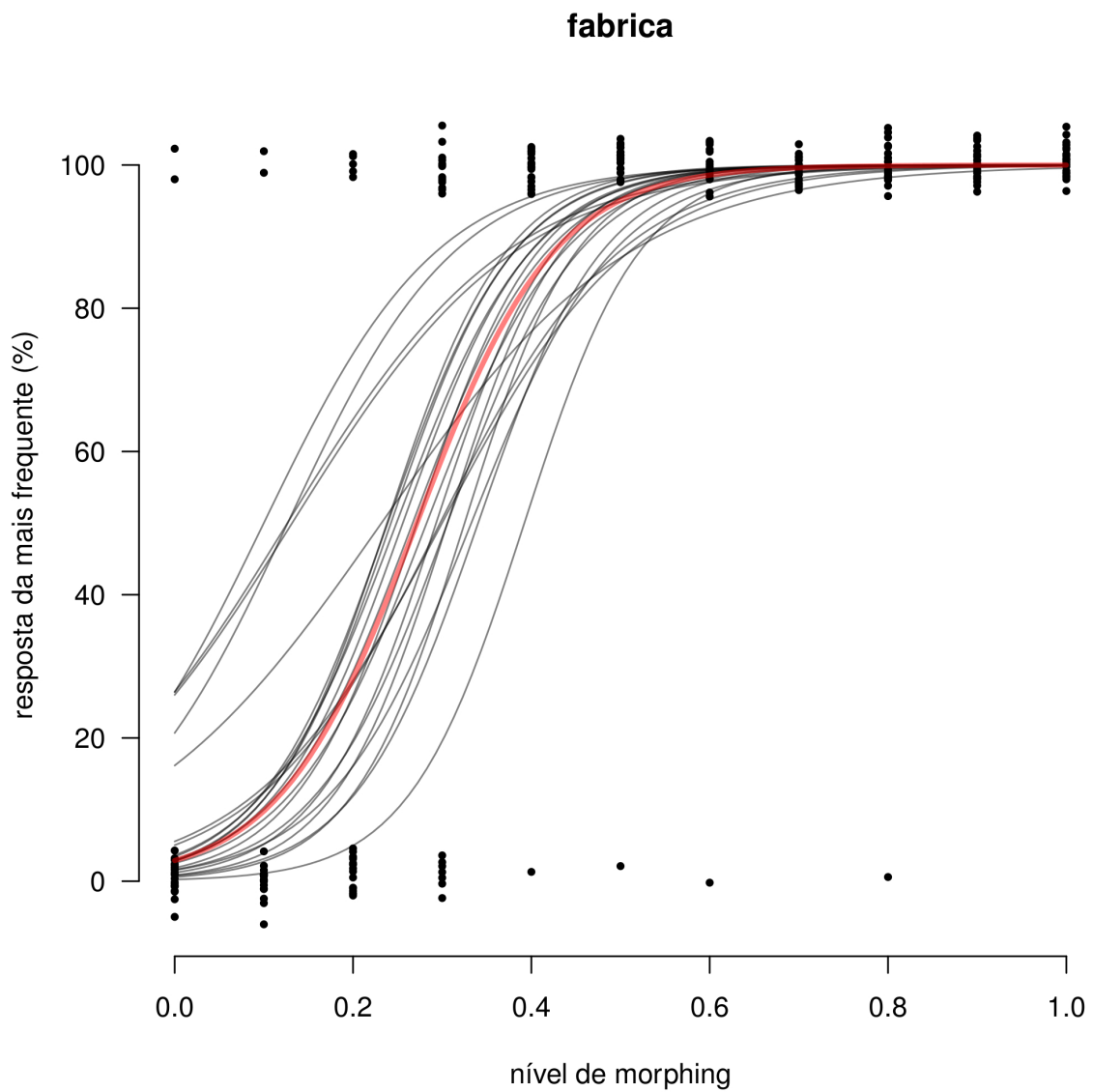


FIGURA 39 - Curva psicométrica do conjunto de estímulos *fabrica-fábrica*

Eixo vertical: 0% = *fabrica*; 100% = *fábrica*

Eixo horizontal: 0.0 = *fabrica*; 1.0 = *fábrica*

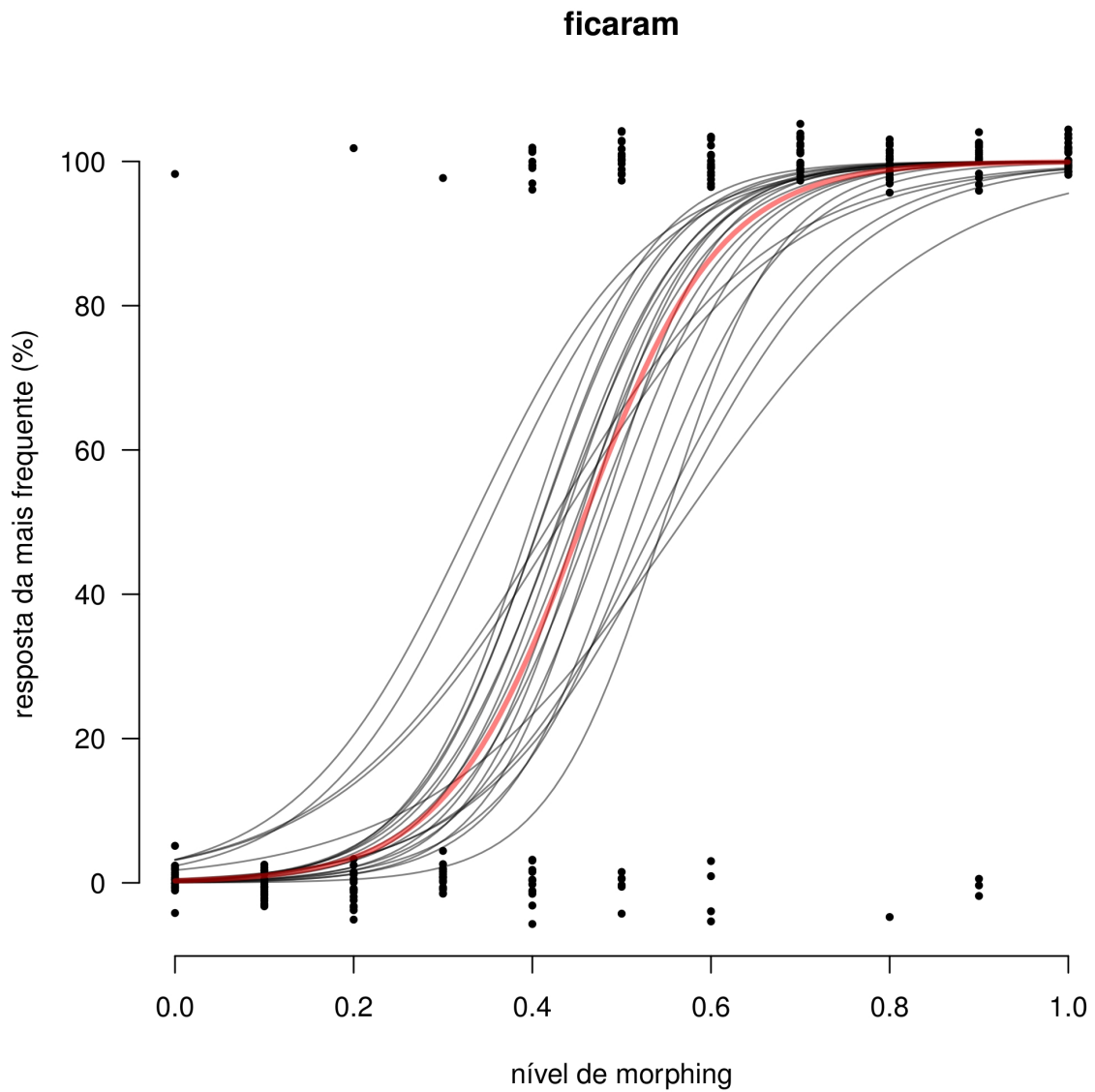


FIGURA 40 - Curva psicométrica do conjunto de estímulos *ficaram-ficarão*

Eixo vertical: 0% = *ficarão*; 100% = *ficaram*

Eixo horizontal: 0.0 = *ficarão*; 1.0 = *ficaram*

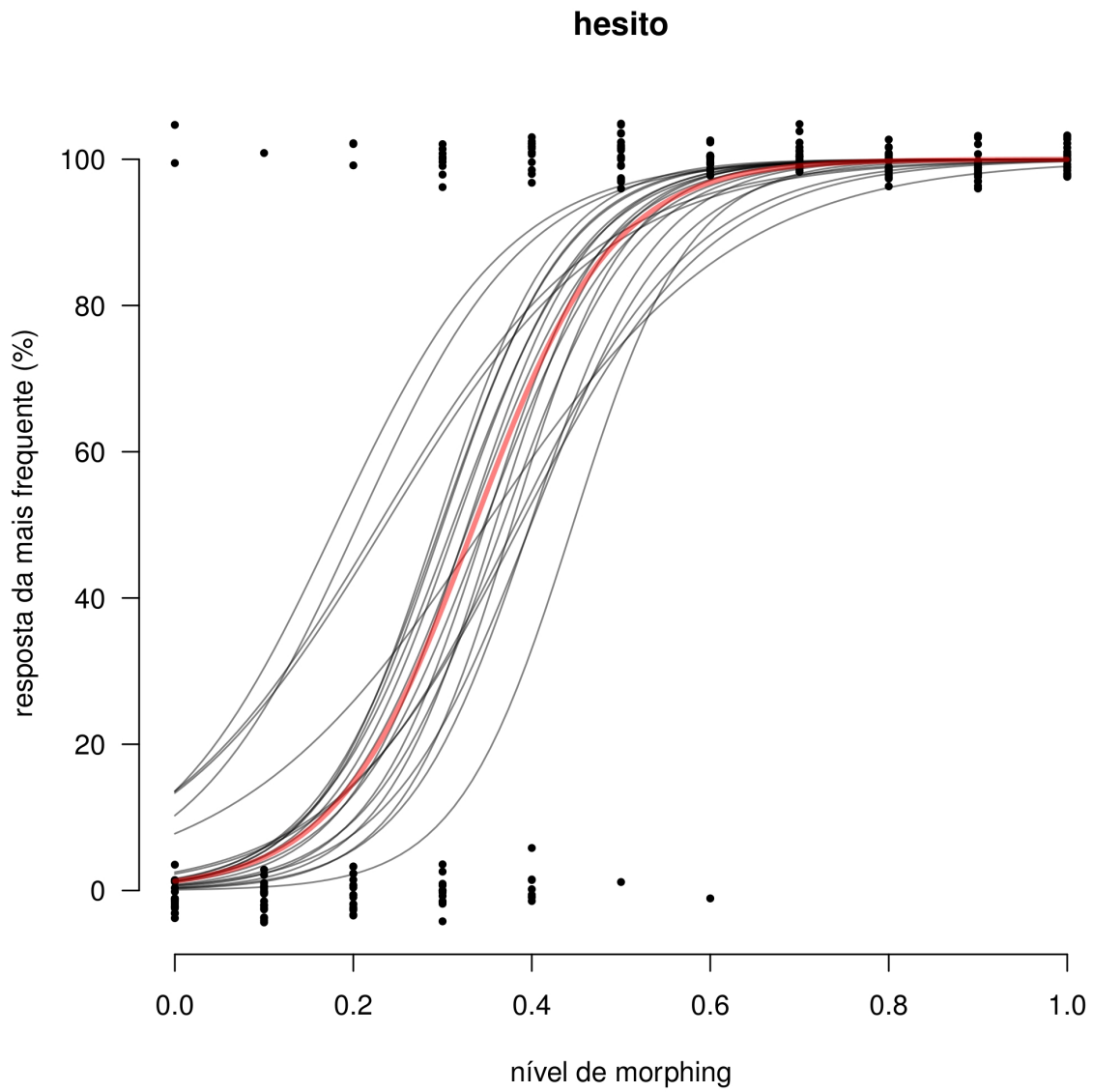


FIGURA 41 - Curva psicométrica do conjunto de estímulos *hesito-êxito*

Eixo vertical: 0% = *hesito*; 100% = *êxito*

Eixo horizontal: 0.0 = *hesito*; 1.0 = *êxito*

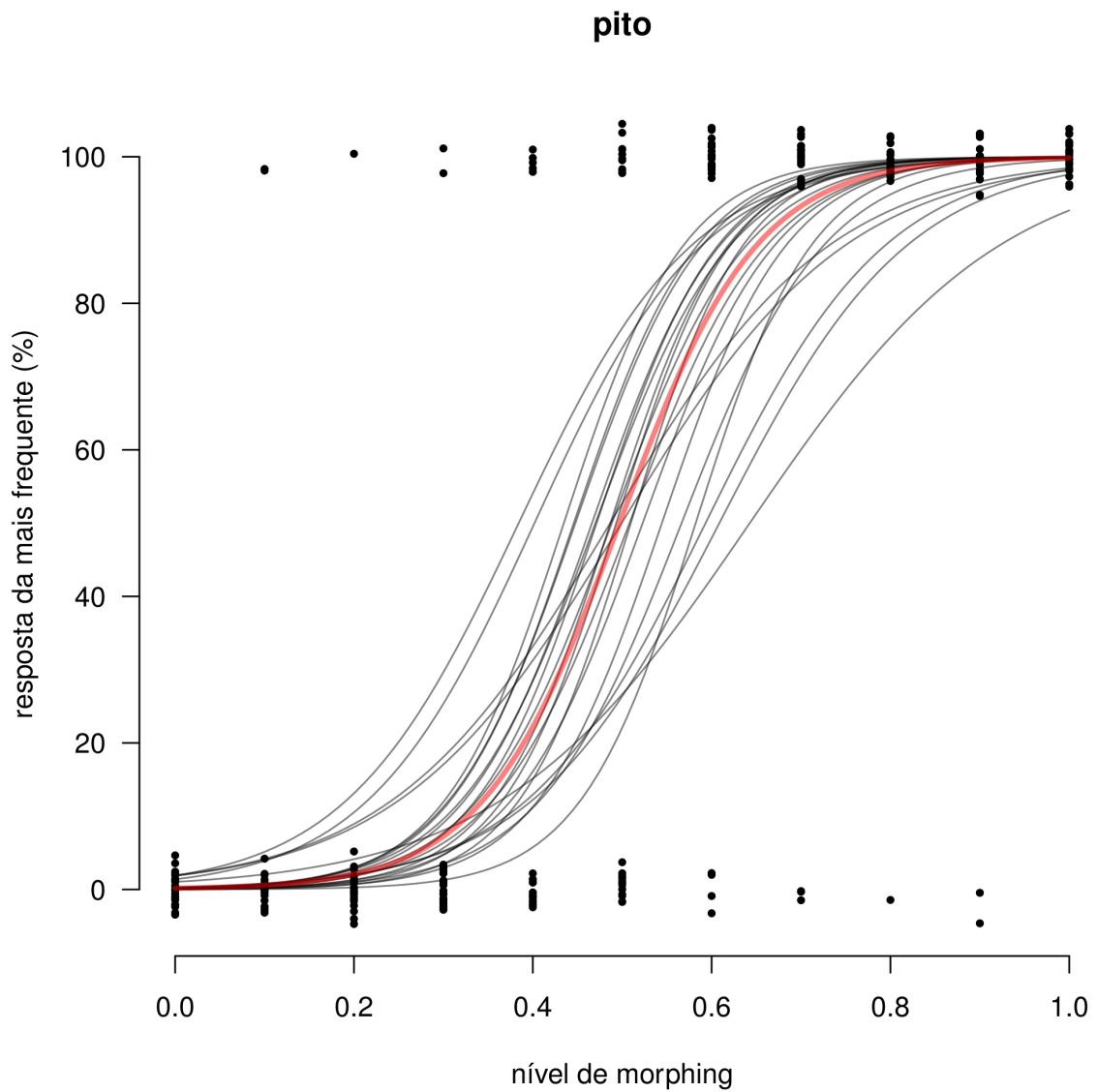


FIGURA 42 - Curva psicométrica do conjunto de estímulos *pito-pitu*

Eixo vertical: 0% = *pitu*; 100% = *pito*

Eixo horizontal: 0.0 = *pitu*; 1.0 = *pito*

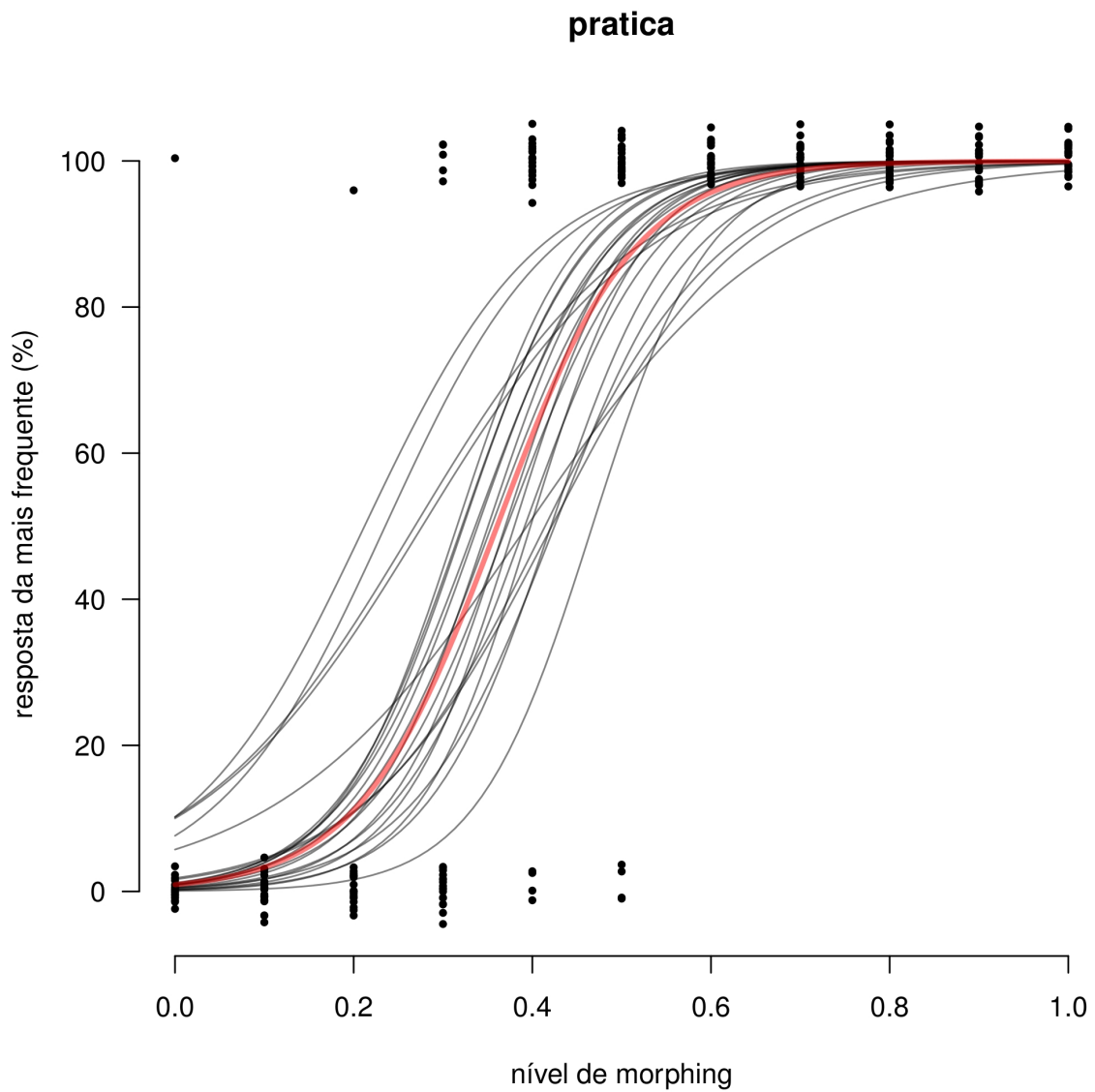


FIGURA 43 - Curva psicométrica do conjunto de estímulos *pratica-prática*

Eixo vertical: 0% = *pratica*; 100% = *prática*

Eixo horizontal: 0.0 = *pratica*; 1.0 = *prática*

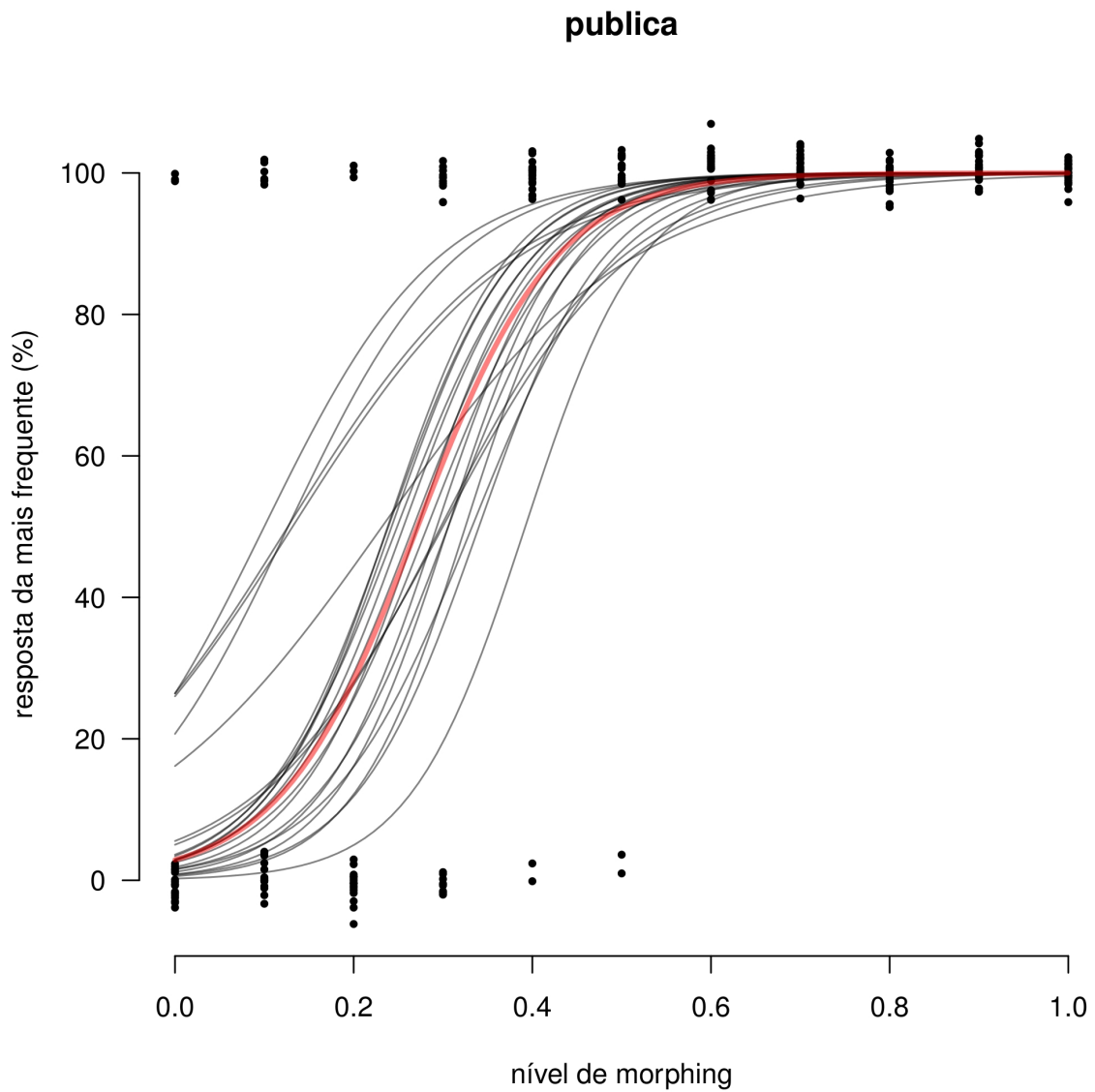


FIGURA 44 - Curva psicométrica do conjunto de estímulos *publica-pública*

Eixo vertical: 0% = *publica*; 100% = *pública*

Eixo horizontal: 0.0 = *publica*; 1.0 = *pública*

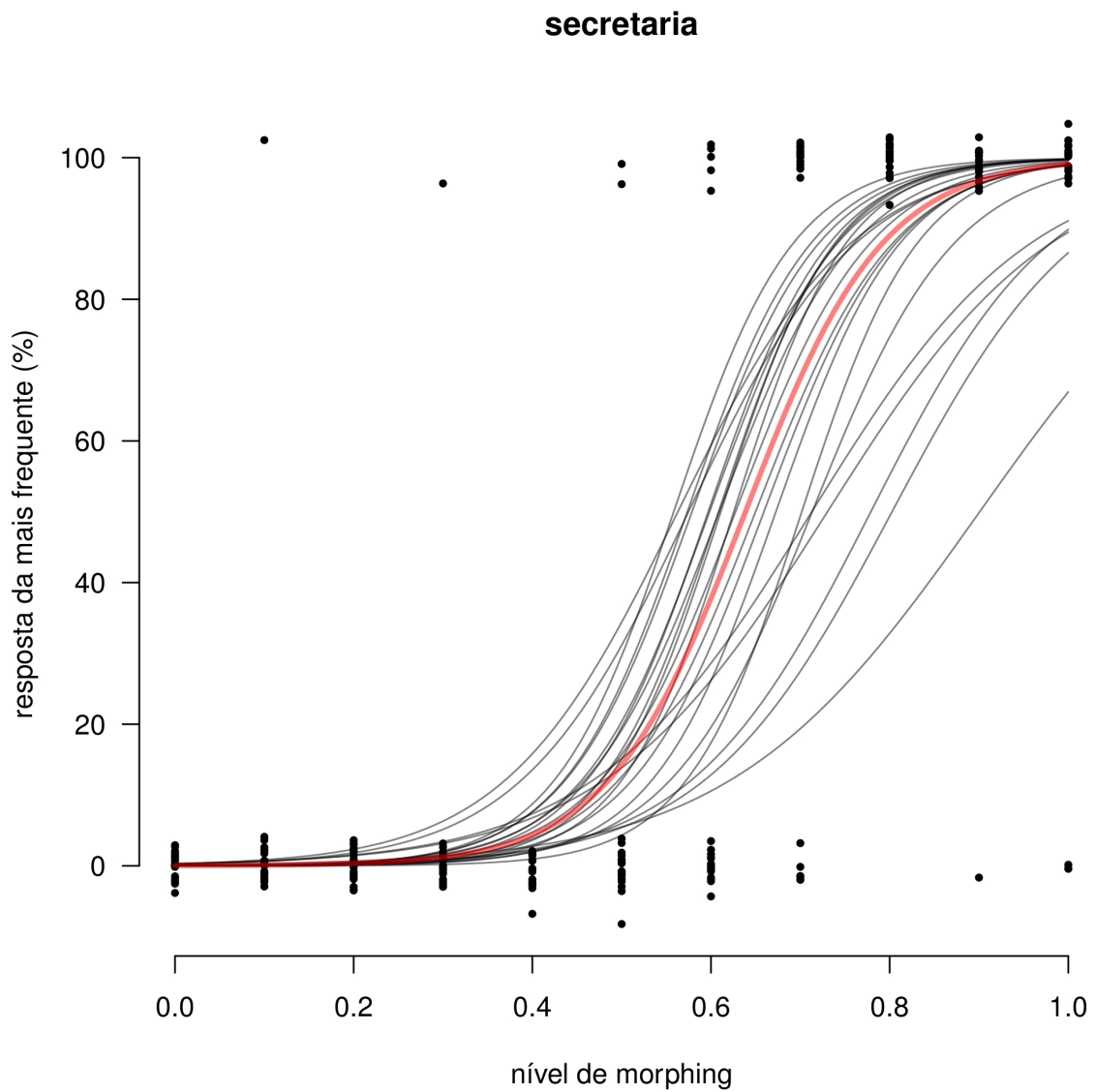


FIGURA 45 - Curva psicométrica do conjunto de estímulos *secretária-secretaria*

Eixo vertical: 0% = *secretária*; 100% = *secretaria*

Eixo horizontal: 0.0 = *secretária*; 1.0 = *secretaria*

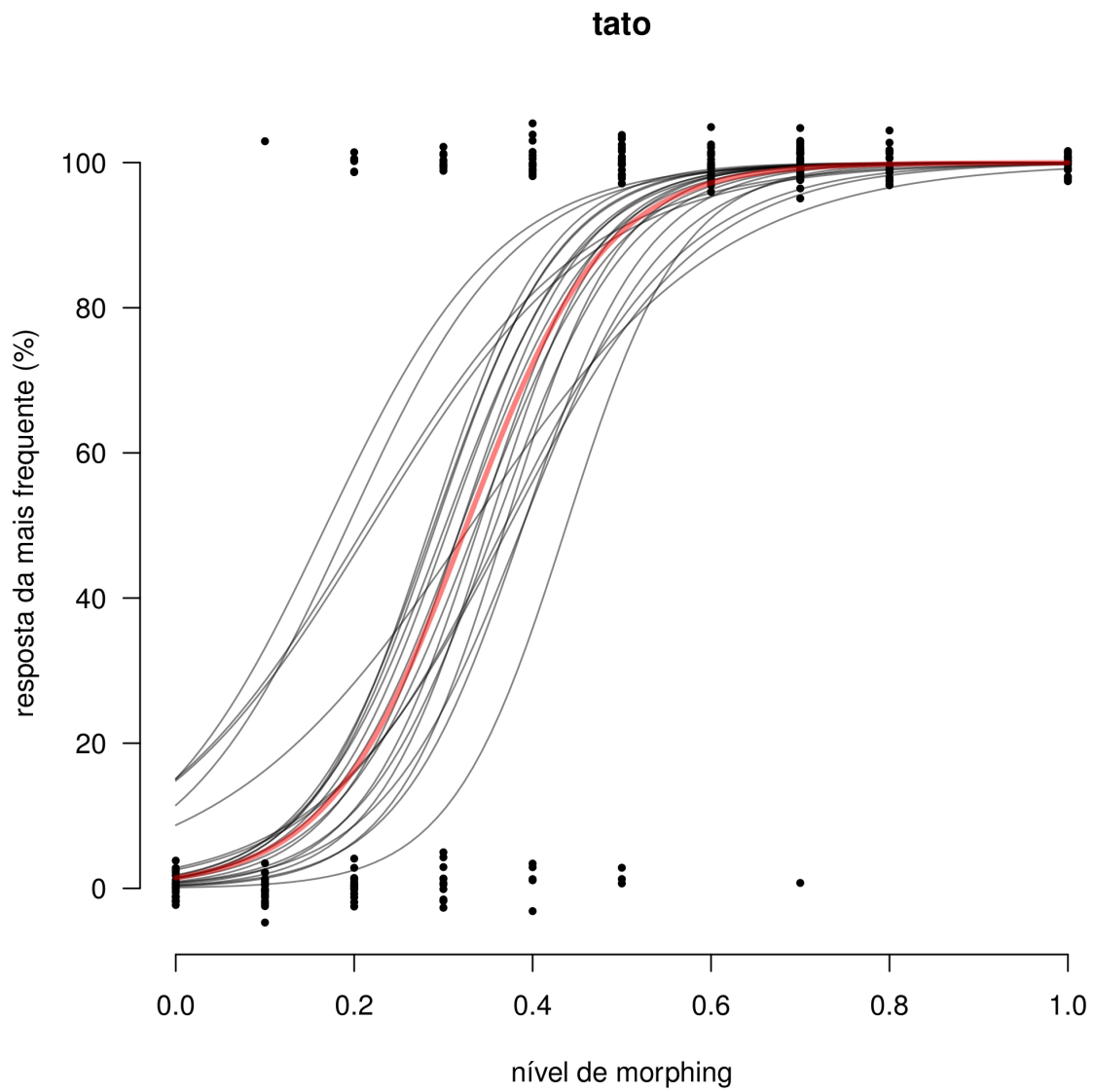


FIGURA 46 - Curva psicométrica do conjunto de estímulos *tato-tatu*

Eixo vertical: 0% = *tatu*; 100% = *tato*

Eixo horizontal: 0.0 = *tatu*; 1.0 = *tato*

## Referências

- ABERCROMBIE, D. *Elements of general phonetics*. Edinburgh: Edinburgh University Press, 1968.
- ALBANO, E. O "tear encantado", a complexidade, a tecnologia digital e o pensamento interdisciplinar sobre a linguagem. *Remate de Males*, v. 29, p. 41-57, 2009.
- ALBANO, E. *O gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro*. Campinas: Mercado de Letras, 2001.
- ALVES, I. M. *Neologismo: criação lexical*. 2. ed. São Paulo: Ática, 1994.
- ARANTES, P. *Integrando produção e percepção de proeminências secundárias numa abordagem dinâmica do ritmo da fala*. Tese (Doutorado) - Campinas, Universidade Estadual de Campinas, 2010a.
- ARANTES, P. *Intensity*. Script para Praat. 2008. Disponível em: <code.google.com/p/praat-tools/>.
- ARANTES, P. *Monofy*. Script para Praat. 2010b. Disponível em: <code.google.com/p/praat-tools/>.
- ARANTES, P. *Slicer*. Script para Praat. 2010c. Disponível em: <code.google.com/p/praat-tools/>.
- ARAÚJO, G. *O acento em português: abordagens fonológicas*. São Paulo: Parábola, 2007.
- ARBIB, M. Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain and Language*, v. 112, p. 12-24, 2010.
- AUDACITY Team. Audacity 1.2.6. 2008. Disponível em: <audacity.sourceforge.net/>. Acesso: 01/05/2010.
- BAAYEN, R. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge: Cambridge University Press, 2008.
- BAAYEN, R.; MCQUEEN, J.; DIJKSTRA, T.; SCHREUDER, R. Frequency effects in regular inflectional morphology: revisiting Dutch plurals. In: BAAYEN, R.; SCHREUDER, R. *Morphological Structure in Language Processing*. New York: Mouton de Gruyter, 2003.
- BARBOSA, P. *Incursões em torno do ritmo da fala*. Campinas: Pontes, 2006.
- BARLOW, M.; KEMMER, S. (Ed.). *Usage-based models of language*. Stanford: CSLI, 2000.
- BATES, D. lme4: Mixed-effects modeling with R. Springer, 2010. Disponível em: lme4.r-forge.r-project.org/IMMwR/lrgprt.pdf. Acesso em 03/11/2012.

- BATES, E.; APPELBAUM, M.; ALLARD, L. Statistical constraints on the use of single cases in neuropsychological research. *Brain and Language*, v. 40, n. 3, p. 295-329, April 1991.
- BECKNER, C. et al. Language is a complex adaptative system: position paper. In: ELLIS, N.; LARSEN-FREEMAN, D. (Ed.). *Language as a Complex Adaptive System*. Oxford: Blackwell, 2009. p. 1-26.
- BELL, A. Accent placement and perception of prominence in rhythmic structures. In: HYMAN, L. (Ed.) *Studies in Stress and Accent*. Los Angeles: University of Southern California, 1977. p. 1-13.
- BISOL, L. (Org.) *Introdução aos estudos de fonologia no português*. 2. ed. Porto Alegre: EDIPUCRS, 1999.
- BISOL, L. O acento e o pé métrico binário. *Cadernos de Estudos Lingüísticos*, v. 22, p. 69-80, 1992.
- BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer*. Versão 5.0.44. Amsterdam: University of Amsterdam, 2008. Disponível em: <[http://www.fon.hum.uva.nl/praat/download\\_win.html](http://www.fon.hum.uva.nl/praat/download_win.html)>. Acesso em: 30 mar. 2009.
- BREGMAN, A. S. *Auditory scene analysis: the perceptual organization of sound*. Cambridge: MIT Press, 1994.
- BROCCIAS, C. Cognitive approaches to grammar. In: KRISTIENSEN, G. *Cognitive linguistics: current applications and future perspectives*. New York: Mouton de Gruyter, 2006. p. 81-115.
- BROWMAN, C.; GOLDSTEIN, L. Towards an articulatory phonology. *Phonology Yearbook*, 3, 219-252, 1986.
- BRUEL-JUNGERMAN, E.; DAVIS, S.; LAROCHE, S. Brain Plasticity Mechanisms and Memory: A Party of Four. *Neuroscientist*, v. 13, p. 492-505, 2007.
- BYBEE, J. *Language, Usage and Cognition*. Cambridge: Cambridge University Press, 2010.
- BYBEE, J. *Morphology: a study of the relation between meaning and form*. Amsterdam: John Benjamins, 1985.
- BYBEE, J. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001.
- BYBEE, J. Regular Morphology and the Lexicon. *Language and Cognitive Processes*, v. 10, n. 5, p. 425-455, 1995.
- BYBEE, J. Sequentiality as the basis of constituent structure. In: GIVÓN, T.; MALLE, B. (Ed.). *The evolution of language out of pre-language*. Amsterdam: John Benjamins, 2002a. p. 109-132.

BYBEE, J. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, v. 14, p. 261-290, 2002b.

BYBEE, J.; HOPPER, P. (Ed.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001.

BYNON, T. *Historical linguistics*. Cambridge: Cambridge University Press, 1986.

BYRD, D.; CHOI, S. At the juncture of prosody, phonology, and phonetics: The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory Phonology 10*. Berlin: Mouton de Gruyter, 2010.

CAGLIARI, L. C. *Acento em português*. Campinas, SP: Edição do autor, 1999.

CAIRNS, H. *The acquisition of language*. Austin: PRO-ED, 1986.

CÂMARA JUNIOR, J. *Estrutura da língua portuguesa*. 2. ed. Petrópolis: Vozes, 1970.

CANTONI, M.; CRISTÓFARO, T. Verbal Stress Assignment in Brazilian Portuguese and the Prosodic Interpretation of Segmental Sequences. In: CONFERENCE ON SPEECH PROSODY, 4th., 2008, Campinas. *Proceedings...* Campinas: IEL, 2008. p. 587-590.

CHOMSKY, N. *The minimalist program*. Cambridge, MA: MIT, 1995.

CHOMSKY, N. Three models for the description of language. *IRE Transactions on Information Theory*, v. 2, n. 3, p. 113-124, 1956.

CHOMSKY, N.; HALLE, M. *The sound pattern of English*. New York: Harper & Row, 1968.

CHOMSKY, N. *Language and mind*. Ed. amp. New York: 1972.

CLEMENTS, N. The geometry of phonological features. *Phonology Yearbook*, v. 2, p. 225-252, 1985.

COLLISCHONN, G. Acento secundário em português. *Letras de Hoje*, v. 29, n. 4, p. 43-53, 1994.

CONNINE, C.; TITONE, D.; WANG, J. Auditory word recognition: Extrinsic and intrinsic effects of word frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, v. 19, n. 1, p. 81-94, Jan 1993.

CONSONI, F. *O acento lexical como pista para o reconhecimento de palavras: uma análise experimental em palavras segmentadas da língua portuguesa*. 2006. Dissertação (Mestrado) – Universidade de São Paulo, São Paulo.

CRAWLEY, M. *The R Book*. Chichester: John Wiley & Sons, 2007.

CRISTÓFARO-SILVA, T.; GOMES, C. Representações múltiplas e organização do

componente linguístico. *Fórum Linguístico*, v. 4, p. 147-177, 2007.

CROFT, W.; CRUSE, D. *Cognitive linguistics*. Cambridge: Cambridge University Press, 2004.

CRUTTENDEN, A. *Intonation*. 2nd. ed. Cambridge: Cambridge University Press, 1997.

CRYSTAL, D. *A dictionary of linguistics and phonetics*. 3rd ed. Oxford: Blackwell, 1995.

CUMMINS, F.; PORT, R. Rhythmic constraints on stress timing in English. *Journal of Phonetics*, v. 26, n. 2, p. 145-171, 1997.

CUTLER, A. Lexical stress. In: PISONI, D. B.; REMEZ, R. E. *Handbook of Speech Perception*. Oxford: Blackwell, 2005. p. 264-289.

CUTLER, A. Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics*, v. 20, n. 1, p. 55-60, 1976.

DOOLING, J. Rhythm and syntax in sentence perception. *Journal of Verbal Learning and Verbal Behavior*, v. 13, n. 3, p. 255-264, June 1974.

DUNN, M.; GREENHILL, S.; LEVINSON, S.; GRAY, R. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, v. 473, n. 7345, p. 79-82, May 2011.

ELLIS, N.; LARSEN-FREEMAN, D. (Ed.). *Language as a Complex Adaptive System*. Oxford: Blackwell, 2009.

ELMAN, J. et. al. *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press, 1996.

FANT, G. *Acoustic theory of speech production: with calculations based on X-Ray studies of Russian articulations*. 2. ed. The Hague: Mouton, 1970 [1960].

FANT, G. The source filter concept in voice production. *Speech, Music and Hearing - Quarterly Progress and Status Report*, v. 22, p. 21-37, 1981.

FARIA, E. *Fonética histórica do latim*. 2. ed. Rio de Janeiro: Acadêmica, 1970.

FARIA, E. *Gramática da língua latina*. 2. ed. rev. aum. Brasília: FAE, 1995.

FARRELL, P. Spanish stress: a cognitive Analysis. *Hispanic Linguistics*, v. 4, n. 1, p. 21-56, Spring 1990.

FERNANDES, N. *Contribuição para uma análise instrumental da acentuação e intonação do Português*. Dissertação (Mestrado) - São Paulo, Universidade Estadual de São Paulo, 1977.

FERREIRA NETO, W. O acento na língua portuguesa. In: ARAÚJO, G. *O acento em português: abordagens fonológicas*. São Paulo: Parábola, 2007. p. 21-36.

FODOR, J. *The modularity of mind: an essay on faculty psychology*. 9th ed. Cambridge, MA: The MIT Press, 1995.

FORSTER, K. I.; FORSTER, J. C. DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, v. 35, p. 116-124, 2003.

FUCHS, S.; WEIRICH, M.; PAPE, D.; PERRIER, P. (Ed.). *Speech Planning and Dynamics*. Frankfurt: Peter Lang, 2012.

GAFOS, A. A Grammar of Gestural Coordination. *Natural Language and Linguistic Theory*, v. 20, p. 269-337, 2002.

GEERAERTS, D. *Cognitive linguistics: basic readings*. Berlin; New York: Mouton de Gruyter, 2006.

GEERAERTS, D. *Theories of lexical semantics*. Oxford University Press, 2010.

GEERAERTS, D.; CUYCKENS, H. Introducing Cognitive Linguistics. In: \_\_\_\_\_ (Ed.). *The Oxford handbook of cognitive linguistics*. New York: Oxford University Press, 2007. p. 3-21.

GELDER, T. van; PORT, R. It's about time: An overview of the dynamical approach to cognition. In: PORT, R.; GELDER, T. van (Ed.). *Mind as Motion: Explorations in the Dynamics of Cognition*. MIT: 1995. p. 1-43.

GOLDINGER, S. On the nature of perceptual adjustments to voice. In: JOHNSON, K.; MULLENIX, J. (Ed.) *Talker variability in speech processing*. San Diego: Academic Press, 1997. p. 33-66.

GOLDSMITH, J. *Autosegmental phonology*. Ph.D. Dissertation - Bloomington, MIT / Indiana University Linguistics Club, 1976.

GOLDSTONE, R.; HENDRICKSON, A. Categorical perception: Advanced Review. *WIREs Cognitive Science*, v. 1, p. 69-78, Dec 2009. Disponível em: <[wires.wiley.com/cogsci](http://wires.wiley.com/cogsci)>. Acesso: 20/03/2012.

GREENBERG, J. *Genetic Linguistics: Essays on theory and method*. Edited by William Croft. New York: Oxford University Press, 2005a.

GREENBERG, J. *Language universals: with special reference to feature hierarchies*. Berlin; New York: Mouton de Gruyter, 2005b.

HADAR, U. Speech-related body movement in aphasia: period analysis of upper arms and head movement. *Brain and Language*, v. 41, p. 339-366, 1991.

HALLE, M.; VERGNAUD, J.-R. Three-dimensional phonology. *Journal of Linguistic*

*Research*, v. 1, p. 83-105, 1987.

HANDEL, S. *An introduction to the perception of auditory events*. Cambridge, MA: MIT, 1993.

HAUSER, M.; CHOMSKY, N.; FITCH, W. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science*, v. 298, p. 1569-1579, Nov. 2002.

HAWKINS, S. Roles and representations of systematic fine phonetic detail in speech understanding. *Journal of Phonetics*, v. 31, p. 373-405, 2003.

HAY, J.; PIERREHUMBERT, J.; BECKMAN, M. Speech perception, well-formedness and the statistics of the lexicon. *Papers in Laboratory Phonology VI*, p. 58-74, 2004.

HAYES, B. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press, 1995.

HEBB, D. *The organization of behavior: A neuropsychological theory*. New York: John Wiley and Sons, 1949.

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

HINTZMAN, D. "Schema Abstraction" in a Multiple-Trace Memory Model. *Psychological Review*, v. 93, n. 4, p. 411-428, 1986.

HOGG, R.; MCCULLY, C. *Metrical Phonology: a coursebook*. Cambridge: Cambridge University Press, 1987.

HYMAN, L. On the nature of linguistic stress. In: \_\_\_\_\_. (Ed.) *Studies in Stress and Accent*. Los Angeles: University of Southern California, 1977. p. 37-82.

HYMAN, L. Word-prosodic typology. *Phonology*, 23, p. 225-257, 2006.

JAKOBSON, R.; FANT, G.; HALLE, M. *Preliminaries to speech analysis*. MIT, 1952.

JAKOBSON, R.; HALLE, M. *Fundamentals of Language*. 2. ed. Paris: Mouton, 1971.

JOHNSON, K. The auditory/perceptual basis for segmentation. *OSU Working Papers in Linguistics*, v. 50, p. 101-113, 1997.

JOHNSON, K.; MULLENIX, J. Complex representations used in speech processing: Overview of the book. In: JOHNSON, K.; MULLENIX, J. (Ed.). *Talker variability in speech processing*. San Diego: Academic Press, 1997. p. 1-8.

JONG, K. The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*. v. 97, n. 1, Jan. 1995.

- KELSO, J. *Dynamic Patterns: The Self Organization of Brain and Behavior*. Cambridge, MA: MIT Press, 1995.
- KEMPS, R.; ERNESTUS, M.; SCHREUDER, R.; BAAYEN, R. Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory Cognition*, v. 33, n. 3, p. 430-446, 2005.
- KENSTOWICZ, M. *Phonology in Generative Grammar*. Oxford: John Wiley and Sons, 1993.
- KENT, R.; READ, C. *The acoustic analysis of speech*. San Diego: Singular, 1992.
- KRISHNAMURTHY, A. *Study of vocal fold vibration and the glottal sound source using synchronized speech, eletroglottograpy and ultra-high speed laryngeal films*. Ph.D. Dissertation – University of Florida, 1983.
- LABOISSIÈRE, R.; ALMEIDA, L.; YEHIA, H.; CRISTÓFARO-SILVA, T. On-going mixing of vowel categories in verbal morphology in Brazilian Portuguese: The importance of fine phonetic detail. *2nd Workshop on Sound Change*, Seon, Germany, 2012.
- LADD, D. R. *Intonational Phonology*. 2nd. ed. Cambridge: Cambridge University Press, 2008.
- LADEFOGED, P. *Elements of Acoustic Phonetics*. Chicago: University Press, 1996.
- LADEFOGED, P. *A course in phonetics*. 2nd ed. Fort Worth: Harcourt Brace College, 1982.
- LADEFOGED, P.; MADDIESON, I. *The Sounds of the World's Languages*. Oxford: Blackwell, 1996.
- LADEFOGED, P.; MCKINNEY, N. Loudness, sound presure level and subglottal pressure in speech. *The Journal of the Acoustical Society of America*, v. 35, n. 4, p. 454-460, April 1963.
- LAKOFF, G. *Women, Fire, and Dangerous Things: what categories reveal about the mind*. Chicago: University of Chicago Press, 1987.
- LANGACKER, R. A Dynamic Usage-Based Model. In: BARLOW, M.; KEMMER, S. (Ed.). *Usage-based models of language*. Stanford: CSLI, 2000. p. 1-63.
- LANGACKER, R. *Foundations of cognitive grammar*. Stanford: Stanford University, 1987. 2 v.
- LANGACKER, R. *Grammar and conceptualization*. Berlin; New York: Mouton de Gruyter, 1999.
- LAVER, J. *Principles of phonetics*. Cambridge: Cambridge University Press, 1994.
- LEE, S. H. Fonologia lexical do português. *Cadernos de Estudos Linguísticos*, v. 23, p. 103-120, jul./dez. 1992.

LEE, S. H. *Morfologia e fonologia lexical do português do Brasil*. Tese (Doutorado) - Campinas, Universidade Estadual de Campinas, Instituto de Estudos da Linguagem, 1995.

LEE, S. H. O acento primário no português: uma análise unificada na teoria da otimalidade. In: ARAÚJO, G. *O acento em português: abordagens fonológicas*. São Paulo: Parábola, 2007. p. 121-144.

LEITE, C. *Sequências de (oclusiva alveolar + sibilante alveolar) como um padrão inovador no português de Belo Horizonte*. 2006. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte.

LEITE, Y. *Portuguese stress and related rules*. Tese (Doutorado) - University of Texas at Austin, 1974.

LIBERMAN, A. M.; HARRIS, K. S.; HOFFMAN, H. S.; GRIFFITH, B. C. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, v. 54, p. 358-368, 1957.

LIBERMAN, M.; PRINCE, A. On stress and linguistic rhythm. *Linguistic Inquiry*, v. 8, p. 249-336, 1977.

LIBREOFFICE. v. 3.3.0. The Document Foundation, 2010. Disponível em: <<http://www.libreoffice.org>>. Acesso: 27/12/2010.

LIEBENTHAL, E.; BINDER, J.; SPITZER, S.; POSSING, E.; MEDLE, D. Neural Substrates of Phonemic Perception. *Cerebral Cortex*, v. 15, p. 1621-1631, October 2005.

MACWHINNEY, B. (Ed.) *The emergence of language*. Mahwah: Lawrence Erlbaum, 1999.

MAJOR, R. Stress and rhythm in Brazilian Portuguese. *Language*, v. 61, n. 2, p. 259-282, 1985.

MARASEK, K. Glottal correlates of the word stress and the tense/lax opposition in German. *Proceedings of the ICSLP*, University of Stuttgart, Stuttgart, v. 96, p. 1573-1576, 1996.

MATLIN, M. *Cognition*. 3rd ed. Hacourt Brace, 1994.

MCCARTHY, J.; PRINCE, A. Prosodic Morphology. In: GOLDSMITH, J. (Ed.). *Phonological theory: The essential readings*. Oxford: Basil Blackwell, 1999.

MEDIN, D.; SCHAFFER, M. Context Theory of Classification Learning. *Psychological Review*, v. 85, n. 3, p. 207-238, 1978.

MEIRELES, A. *Reestruturações rítmicas da fala no português brasileiro*. 2007. Tese (Doutorado em Lingüística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas.

MOATES, D.; SCHUMACHER, G. *An introduction to cognitive psychology*. Belmont:

Wadsworth, 1980.

MOOSHAMMER C.; FUCHS, S. Stress distinction in German: Simulating kinematic parameters of tongue tip gestures. *Journal of Phonetics*, v. 30, p. 337-355, 2002.

MORAES, J. A. Intonation in Brazilian Portuguese. In: HIRST, D.; DI CRISTO, A. *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 1998.

NESPOR, M.; VOGEL, I. *Prosodic phonology*. Dordrecht: Foris, 1986.

NIELSEN, K. Effect of lexical frequency and neighborhood density on audiovisual spoken recognition. *ICPhS XVII*, 2011.

NOSOFSKY, R. Attention, Similarity, and the Identification-Categorization Relationship. *Journal of Experimental Psychology*, v. 115, n. 1, p. 39-57, 1986.

OLIVEIRA, D. Gradualidade e frequência: contribuições do Modelo de Exemplar e da Fonologia de Uso ao estudo da variação sonora nas seqüências de (sibilante + africada alveopalatal). Trabalho de conclusão da disciplina LIG917. FALE, UFMG. 2003. ms.

ORTEGA-LLEBARIA, M.; PRIETO, P. Disentangling stress from accent in Spanish: production patterns of stress contrast in deaccented syllables. In: PRIETO, P.; MASCARÓ, J.; SOLÉ, M. J. (Ed.). *Segmental and Prosodic Issues in Romance Phonology*. John Benjamins: Amsterdam/ Philadelphia, 2007. p. 155-175.

PAGEL, M.; ATKINSON, Q.; MEADE, A. Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature*, v. 449, p. 717-720, 2007.

PALMER, G. *Toward a theory of cultural linguistics*. Austin: University of Texas Press, 1999.

PIERREHUMBERT, J. Exemplar dynamics: Word frequency, lenition and contrast. In: BYBEE, J.; HOPPER, P. (Ed.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001a. p. 137-157.

PIERREHUMBERT, J. Stochastic phonology. *GLOT*, v. 5, n. 6, p. 1-13, 2001b.

PIERREHUMBERT, J. Probabilistic Phonology: discrimination and robustness. In: BOD, R.; HAY, J.; JANNEDY, S. (Ed.). *Critical Introduction to Phonology*. Cambridge, MA: MIT, 2003. p. 177-228.

PIERREHUMBERT, J. *The Phonology and Phonetics of English Intonation*. Ph.D. Dissertation - MIT, Cambridge, MA, 1980.

PIERREHUMBERT, J.; BECKMAN, M. *Japanese Tone Structure*. Cambridge, MA: MIT Press, 1988.

PIERREHUMBERT, J.; BECKMAN, M.; LADD, R. Conceptual foundations of phonology as

a laboratory science. In: BURTON-ROBERTS, N.; CARR, P.; DOCHERTY, G. (Ed.). *Phonological Knowledge*. Oxford: Oxford University Press, 2000. p. 273-303.

PINKER, S. *The language instinct*. New York: William Morrow, 1994.

PHILLIPS, B. Lexical diffusion, lexical frequency, and lexical analysis. In: BYBEE, J.; HOPPER, P. (Ed.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001. p. 123-136.

PHILLIPS, B. *Word frequency and lexical diffusion*. New York: Palgrave Macmillan, 2006.

POPPER, K. *The logic of scientific discovery*. New York: Harper TorchBooks, 1968.

PORT, R. How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology*, v. 25, p. 143-170, 2007.

PRINCE, A.; SMOLENSKY, P. Optimality Theory: Constraint Interaction in Generative Grammar. *Rudgers Optimality Archive (ROA)*, 2002. Disponível em: <<http://roa.rutgers.edu/>>. Acesso: 25/08/2008.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. v. 2.13.0, Vienna – Áustria, 2011. Disponível em: <<http://www.r-project.org/>>. Acesso: 26/04/2011.

REPP, B. Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, v. 92, p. 81-110, 1982.

REPP, B. Categorical perception: Issues, methods, findings. In: LASS, N. (Ed.). *Speech and language: Advances in basic research and practice*, v. 10, p. 243-335, 1984.

ROSCH, E. Principles of categorization. In: ROSCH, E.; LLOYD, B. (Ed.). *Cognition and Categorization*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1978. p. 27-48.

ROSENZWEIG, M.; POSTMAN, L. Frequency of usage and the perception of words. *Science*, v. 127, p. 263-266, 1958.

ROTHENBERG, M. A new inverse-filtering technique for deriving the glottal airflow waveform during voicing. *The Journal of the Acoustical Society of America*, v. 53, n. 6, p. 1632-1645, June 1973.

ROTHENBERG, M.; MASHIE, J. Monitoring Vocal Fold Abduction Through Vocal Fold Contact Area. *Journal of Speech and Hearing Research*, v. 31, p. 338-351, Sep. 1988.

SALTZMAN, E., NAM, H., KRIVOKAPIC, J.; GOLDSTEIN, L. A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Proceedings of Speech Prosody 2008*.

SALTZMAN, E.; MUNHALL, K. A dynamical approach to gestural patterning in speech

production. *Ecological Psychology*, v. 1, n. 4, p. 333-382, 1989.

SAMPSON, G. *The "language instinct" debate*. London; New York: Continuum, 2005.

SCHLICK, M.; CARNAP, R.; POPPER, K. *Coletânea de textos*. São Paulo: 1975.

SELKIRK, E. The role of prosodic categories in English word stress. *Linguistic Inquiry*, v. 11, p. 563-605, 1980.

SILVA, R. V. *O português arcaico: fonologia*. 2. ed. São Paulo: Contexto, 1995.

SILVA, T. *Fonética e fonologia do português: roteiro de estudos e guia de exercícios*. 6. ed. rev. São Paulo: Contexto, 2002

SILVERMAN, D. *A critical introduction to phonology: Of sound, mind, and body*. Continuum International, 2006.

SINHA, C. Biology, Culture and the Emergence and Elaboration of Symbolization. In: SALEEMI, A.; BOHN, O.; GJEDDE, A. (Ed.). *In search of a language for the mind/brain: Can the Multiple Perspectives be Unified?* Aarhus: Aarhus University Press, 2005. p. 311-335.

SINHA, C. Cognitive linguistics, psychology, and cognitive science. In: GEERAERTS, D.; CUYCKENS, H. (Ed.). *The Oxford handbook of cognitive linguistics*. New York: Oxford University Press, 2007. p. 1266-1294.

SINHA, C. *Language and representation: A socio-naturalistic approach to human development*. New York: New York University Press, 1988.

SINHA, C. Signifying Subjects. In: HAUKIOJA, T. (Ed.). *Papers on Language Theory*. Turku: University of Turku, 2006. p. 147-177.

SLUIJTER, A.; HEUVEN, V. van. Spectral balance as an acoustical correlate of linguistic stress. *Journal of the Acoustical Society of America*, v. 100, n. 4, p. 2471-2485, 1996.

SOUZA, R. *A redução de vogais altas pretônicas no português de Belo Horizonte: uma abordagem baseada na gradiência*. 2012. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, Belo Horizonte.

STEELS, L. Modeling the cultural evolution of language. *Physics of Life Reviews*, v. 8, p. 339-356, 2011.

STEVENS, K. *Acoustics Phonetics*. Cambridge, MA: MIT, 2000.

TAYLOR, J. *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Clarendon, 1991.

TAYLOR, J. Polysemy and the lexicon. In: KRISTIANSEN, G. *Cognitive linguistics: current applications and future perspectives*. New York: Mouton de Gruyter, 2006. p. 51-80.

TEYSSIER, P. *História da língua portuguesa*. 2. ed. Tradução de Celso Cunha. São Paulo: Martins Fontes, 2001.

THELEN, E.; SMITH, L. *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press, 1994.

TOMASELLO, M. *The cultural origins of human cognition*. Harvard University Press, 1999.

TURK, A.; SAWUSCH, J. The processing of duration and intensity cues to prominence. *Journal of the Acoustical Society of America*, v. 99, n. 6, p. 3782-3790, 1996.

VAISSIERE, J. Perception of intonation. In: PISONI, D. B.; REMEZ, R. E. *Handbook of Speech Perception*. Oxford: Blackwell, 2005. p. 236-263.

VIEIRA, M. *Automated measures of dysphonias and the phonatory effects of asymmetries in the posterior larynx*. 1997. Ph.D. Thesis – University of Edinburgh, Edinburgh.

VIGOTSKI, L. *A construção do pensamento e da linguagem*. Tradução de Paulo Bezerra. São Paulo: Martins Fontes, 2001.

WETZELS, L. Primary Word Stress in Brazilian Portuguese and the Weight Parameter. *Journal of Portuguese Linguistics*, v. 5, n. 2, p. 9-58, 2006.

WILLIAMS, E. *Do latim ao português: fonologia e morfologia históricas da língua portuguesa*. 3. ed. Rio de Janeiro: Tempo Brasileiro, 1975.

WHALEN, D. H.; LEVITT, A. G. The universality of intrinsic F0 of vowels. *Journal of Phonetics*, v. 23, p. 349-366, 1995.

ZÁGARI, M. R. *Fonologia diacrônica do português*. Juiz de Fora: Universidade Federal de Juiz de Fora, 1988.