UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

O ACENTO PRIMÁRIO EM PSEUDOPALAVRAS:

uma abordagem experimental

ALINE DE LIMA BENEVIDES

VERSÃO CORRIGIDA

SÃO PAULO
2017

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

O ACENTO PRIMÁRIO EM PSEUDOPALAVRAS:

uma abordagem experimental

ALINE DE LIMA BENEVIDES

VERSÃO CORRIGIDA

Dissertação apresentada ao Programa de Pós-Graduação em Linguística do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Mestre em Letras.

Orientador: Prof. Dr. Paulo Chagas de Souza
Coorientadora: Prof^a. Dr^a. Thaïs Cristófaró Silva

De acordo:

A handwritten signature in blue ink, reading "Paulo Chagas de Souza", written over a horizontal line.

SÃO PAULO
2017

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

B456a Benevides, Aline de Lima
O Acento primário em pseudopalavras: uma
abordagem experimental / Aline de Lima Benevides ;
orientador Paulo Chagas de Souza. - São Paulo, 2017.
135 f.

Dissertação (Mestrado)- Faculdade de Filosofia,
Letras e Ciências Humanas da Universidade de São
Paulo. Departamento de Linguística. Área de
concentração: Semiótica e Linguística Geral.

1. Acento. 2. Fonologia de Laboratório. 3. Modelos
Multirrepresentacional. 4. Pseudopalavras. 5.
Português Brasileiro. I. Souza, Paulo Chagas de,
orient. II. Título.

BANCA EXAMINADORA

Prof. Dr. Paulo Chagas de Souza (USP)
Presidente

Prof^a. Dr^a. Thaïs Cristófaro Silva (UFMG)
Vice-Presidente

Prof^a. Dr^a. Raquel Santana Santos (USP)
Membro titular

Prof. Dr. Waldemar Ferreira Netto (USP)
Membro titular

Prof^a. Dr^a. Maria Mendes Cantoni (UFMG)
Membro titular

*Aos meus pais,
Francisco (in memoriam) e Ana,
ao meu amigo e companheiro,
Leonardo,
e aos meus pequenos.*

Agradecimentos

Ao meu pai, Francisco, por me ensinar o valor do conhecimento e por me mostrar o quão forte posso ser - aqui fica minha eterna saudade. À minha mãe, Ana, que desde muito cedo me ensinou a lutar e a buscar os meus sonhos, que apoiou, muitas vezes, sem compreender, as minhas escolhas e caminhos.

Ao meu companheiro e amigo, Leonardo, que acompanhou cada etapa desta pesquisa, que soube lidar com todas as minhas crises, angústias e alegrias provenientes de uma longa jornada de trabalho, que discutiu, rediscutiu e ouviu todas as minhas hipóteses. E antes de mais nada, pelo apoio, companheirismo e amor ao longo dessa caminhada.

Ao meu amigo e orientador, Paulo Chagas de Souza, que despertou em mim o amor pela Fonologia, pela pesquisa e, principalmente, pelo acento, que desde a Iniciação Científica me acompanha, atende aos meus e-mails e mensagens desesperados e com tanta sensibilidade soube compreender todas as minhas dificuldades e desafios em cada uma das etapas de minha formação.

À minha amiga e coorientadora, Thaïs Cristófaró Silva, que me encontrou perdida em alguns congressos e acreditou no meu potencial, sendo a principal fonte de diálogo, de discussões e que, apesar de sua agenda lotada, reservou sempre minutos preciosos para me atender, me receber e discutir comigo cada ponto dos resultados desta pesquisa, minimizando cada quilômetro de distância com seus atendimentos via web.

Aos meus irmãos, Alexandre e Ariane, e aos meus cunhados, Natália e Lucas, que me incentivaram e que me auxiliarem em idas e vindas a inúmeros congressos e seminários. Aos meus primos, Janaina, Jaqueline, Alessandra, Marcelo e Diego, que prontamente, desde a minha Iniciação Científica, aceitaram aos meus convites para colaborar com os mais diferentes e, muitas vezes, longos testes para verificar os desenhos experimentais.

A Eder de Lucca e Leandro Rena, por terem acreditado no meu potencial, e me encorajado, em um momento em que eu nem mesmo sonhava. Aos meus queridos, Ana Carolina, Ana Paula, Andressa Nesi, Camilla Rezende, Fernanda Medeiros, Joice Rodrigues, Júlia Fernandez, Karina Oliveira, Raquel Silva, Rodrigo Godinho e Wellington Silva, pelo apoio, carinho, amizade,

confiança e incentivo recebidos ao longo de toda a minha formação. Agradeço, ainda, a Nívia e a Pepita por tão carinhosamente me receberem em sua casa com tanta alegria e atenção.

A Fernanda Canever, Fátima Baia, Walter Sano e, em especial, a Livia Oushiro, por serem exemplos de pessoas, pesquisadores, colegas, amigos e por sempre estarem dispostos a me ajudar, independente do horário, do dia ou da dificuldade. Agradeço, ainda, por compartilharem comigo inúmeras experiências de viagens, congressos e a real face da vida acadêmica.

Ao colega de mestrado, Bruno Guide, pela parceria na compilação do *corpus*, no desenvolvimento de ferramentas computacionais e por toda paciência nos estudos de programação. Ao Prof. Dr. Leonardo Araújo, por dispor inúmeras vezes do seu tempo para me ajudar a extrair e manipular os dados e pelos valiosos comentários e discussões compartilhadas ao longo do desenvolvimento desta pesquisa.

Aos professores, Ana Müller, Beatriz Raposo, Livia Oushiro, Ronald Beline e Paulo Chagas pelos cursos enriquecedores que pude frequentar; ao Marcos Lopes e ao Marcelo Ferreira, por todo acompanhamento no desenvolvimento de programas e do *corpus*. Em especial, agradeço a Raquel Santana Santos, pelos comentários e contribuições fornecidos no exame de qualificação e na elaboração do desenho experimental desta pesquisa.

Aos membros do GLiC (Grupo de Linguística Computacional da USP), do Grupo de Fonologia Cognitiva e, principalmente, do FONEMOS (Grupo de Estudos de Fonologia e Morfologia da USP), os quais sempre estiveram dispostos a ouvir, contribuir e discutir todas as etapas deste estudo.

Agradeço às professoras da Universidade de Lisboa, Marisa Cruz e Paula Luegi, por me introduzirem na área experimental, mostrando e demonstrando com imensa atenção as metodologias, as técnicas e os cuidados necessários para a realização de experimentos.

A todos os conferencistas, ouvintes e palestrantes que, diretamente ou indiretamente, contribuíram por meio de aulas, palestras e comentários com o andamento desta pesquisa. Em especial, ao Prof. Dr. Seung Haw Lee, que, em um congresso em 2013, plantou a semente que germinou com a realização deste trabalho.

Aos membros da banca, Maria Mendes Cantoni, Raquel Santana Santos e Waldemar Ferreira Netto, e aos suplentes, Clerton Luiz Felix Barbosa, Gabriel Antunes de Araújo e José Sueli Magalhães, por aceitarem o convite de compor a banca deste trabalho, pelos comentários e disponibilidade.

A todos os informantes, amigos, amigos de amigos e alunos do curso de Letras, que com imensa dedicação realizaram o experimento deste estudo, sem os quais este trabalho não teria se realizado.

Ao Robson, à Érica e à Denise que auxiliaram em todos os processos burocráticos com imensa atenção.

Pelo financiamento desta pesquisa, agradeço ao CNPq.

*BEGINNING my studies, the first step pleas'd me so much,
The mere fact, consciousness-these forms-the power of motion,
The least insect or animal—the senses—eyesight—love;
The first step, I say, aw'd me and pleas'd me so much,
I have hardly gone, and hardly wish'd to go, any farther,
But stop and loiter all the time, to sing it in extatic songs.
(Walt Whitman)*

Resumo

Esta dissertação tem por objetivo investigar o papel que a frequência dos tipos silábicos desempenha na atribuição acentual em pseudopalavras produzidas por falantes nativos do Português Brasileiro através de metodologias experimentais. Contrapuseram-se, para isso, dados empíricos, provenientes de um experimento, a hipóteses fonológicas para o acento primário em PB. As propostas analisadas consistiram em: Hipótese Lexical (CÂMARA JR., 1970 [2001]), Hipótese Métrica (BISOL, 1994), Hipótese Morfológica (LEE, 1995) e Hipótese Multirrepresentacional (CANTONI, 2013). Esta pesquisa está norteada nos pressupostos dos Modelos Multirrepresentacionais. Fez-se necessária a compilação de um *corpus* linguístico que permitisse a extração da frequência dos padrões fonológicos requeridos na presente investigação, o qual é denominado de *Corpus* ABG. O experimento, que se subdividiu em dois grupos experimentais, inquiriu a relevância dos tipos silábicos [niw], [V_{oral}.tʃiw], [V_{nasal}.tʃiw], [-or] e [-ar] nos estudos acentuais. O grupo de análise 1, composto pelos tipos silábicos [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], investigou, por meio do conceito de lacuna fonológica, se a frequência dos tipos silábicos impacta na determinação da localização acentual. O grupo de análise 2, composto pelos tipos silábicos [-or] e [-ar], verificou em que medida a frequência dos tipos silábicos interage com processos analógicos, tendo como motivação a frequência de vocábulos fonológicos similares. Os resultados do experimento sugerem que, ao contrário do que preveem as propostas algorítmicas, todos os padrões acentuais (proparoxítono, paroxítono e oxítono) podem emergir - e emergiram. As produções acentuais encontradas no experimento trazem indícios de que a frequência dos tipos silábicos impacta na forma como o acento é atribuído no PB, bem como trazem evidências de que processos analógicos podem interagir com a produtividade na atribuição acentual. Os resultados, portanto, indicam que diferentes fatores, em diferentes níveis de abstração, são responsáveis pela atribuição acentual no PB.

Palavras-chave: acento; fonologia de laboratório; modelos multirrepresentacionais; pseudopalavras; Português Brasileiro.

Abstract

This dissertation aims to investigate the role that the frequency of the syllabic types plays in the stress assignment in pseudowords produced by native speakers of Brazilian Portuguese through experimental methodologies. For this, empirical data, originating from an experiment, were contrasted to phonological hypotheses for the primary stress in BP. The proposals analyzed consisted of: Lexical Hypothesis (CÂMARA JR., 1970 [2001]), Metrical Hypothesis (BISOL, 1994), Morphological Hypothesis (LEE, 1995) and Multirepresentational Hypothesis (CANTONI, 2013). This research is guided on the assumptions of the Multirepresentational Models. A compilation of a linguistic corpus was necessary to allow the extraction of the frequency of the phonological patterns required in the present investigation, which is named Corpus ABG. The experiment, which was subdivided into two experimental groups, inquired about the relevance of the syllabic types [niw], [V_{oral}.tʃiw], [V_{nasal}.tʃiw], [-or] and [-ar] in the stress studies. Analysis group 1, composed of the syllabic types [niw], [V_{oral}.tʃiw] and [V_{nasal}.tʃiw], investigated, through the concept of phonological gap, whether the frequency of the syllabic types has an impact on the determination of the stress location. Analysis group 2, composed of the syllabic types [-or] and [-ar], verified to what extent the frequency of the syllabic types interacts with analogical processes, motivated by the frequency of similar phonological vocables. The results of the experiment demonstrated that, contrary to what the algorithmic proposals foresee, all the stress patterns (proparoxitone, paroxitone and oxytone) can emerge - and emerged. The stress productions found in the experiment indicate that the frequency of the syllabic types impacts on how the stress is assigned in BP, as well as provides evidences that analogical processes can interact with productivity in stress assignment. The results, therefore, demonstrate that different factors, at different levels of abstraction, are responsible for the stress assignment in BP.

Keywords: stress; laboratory phonology; multirepresentational models; pseudowords; Brazilian Portuguese.

Sumário

Lista de Tabelas	xiv
Lista de Quadros	xv
Lista de Figuras	xvii
Introdução	1
1 Pressupostos Teóricos	5
1.1 O acento na perspectiva fonética	5
1.2 Sílabas & Acento	7
1.3 O acento na perspectiva fonológica	10
1.3.1 Hipótese Lexical	10
1.3.2 Hipótese Métrica	13
1.3.3 Hipótese Morfológica	16
1.3.4 Hipótese Multirrepresentacional	19
1.4 Modelos Multirrepresentacionais	25
1.5 Considerações Finais	31
2 Metodologia	33
2.1 Desafios Metodológicos	33
2.1.1 A Ortografia	34
2.1.2 Conhecimento Ortográfico	36
2.1.3 Metaexperimento	40
2.1.4 <i>Corpus</i> ABG	46
2.2 Metodologia	51

2.2.1	Participantes	51
2.2.2	Estímulos	52
2.2.3	Quantidade de Dados	61
2.2.4	Procedimentos	61
2.3	Método de Análise dos Dados	63
2.3.1	Análise Perceptiva	63
2.3.2	Variáveis e Modelos Estatísticos	64
2.4	Considerações Finais	66
3	Resultados e Discussão	68
3.1	Resultados Gerais	69
3.2	Grupo de Análise 1: [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw]	71
3.3	Grupo de Análise 2: [-or] e [-ar]	87
3.4	Proparoxítonas	99
3.5	Informante	103
3.6	Considerações Finais	107
4	Conclusão	109
	Referências Bibliográficas	114
	Anexo A: Resultados do metaexperimento	125
	Anexo B: Resultados estatísticos do metaexperimento	126
	Anexo C: Pré-teste	127
	Anexo D: Gráfico dos modelo de regressão logística no Grupo de Análise 1	130
	Anexo E: Gráfico do modelo de regressão logística com efeitos aleatórios no Grupo de Análise 1	131
	Anexo F: Resultados do modelo no Grupo de Análise 1	132
	Anexo G: Resultados do modelo no Grupo de Análise 2	134

Lista de Tabelas

44table.caption.17

2.2	Frequência de ocorrência e de tipo por <i>corpus</i>	48
2.3	Quantidade de tipos de palavras por categoria acentual.	49
2.4	Distribuição percentual do acento por categoria em Viaro & Guimarães-Filho (2007, p. 31) e no <i>Corpus</i> ABG.	49
2.5	Tonicidade por estrutura silábica final no subABG.	51
2.6	Frequência de tipo das sequências [niw] e [tʃiw] por tonicidade no <i>Corpus</i> ABG.	54
2.7	Pseudopalavras do Grupo de Análise 1.	56
2.8	Frequência de tipo das sequências [-or] e [-ar].	58
2.9	Pseudopalavras do Grupo de Análise 2.	60
3.1	Atribuição acentual nos dados gerais.	70
3.2	Frequência de tipo das sequências [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw] por <i>tonicidade</i> no <i>Corpus</i> ABG e nos resultados do Grupo de Análise 1.	73
3.3	Atribuição acentual no grupo de análise 1 - [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw].	73
3.4	Resultados reestruturados de atribuição acentual em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw] (sem proparoxítonas).	78
3.5	Correlação entre a variável dependente e as variáveis independentes em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw].	78
3.6	Resultados do modelo de regressão logística com efeitos mistos em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw].	81

3.7	Resultado de atribuições acentuais por <i>pseudopalavra</i> em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw] (em ocorrências).	85
3.8	Resultado de atribuição acentual por <i>repetição</i> em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw].	86
3.9	Frequência de tipo das sequências sonoras [-or] e [-ar] por <i>tonicidade</i> no <i>Corpus</i> ABG e os resultados obtidos no Grupo de Análise 2.	88
3.10	Atribuição acentual no grupo de análise [-or] e [-ar].	90
3.11	Resultados reestruturados de atribuição acentual em [-or] e [-ar] (sem proparoxítonas).	93
3.12	Correlação entre a variável dependente e as variáveis independentes em [-or] e [-ar].	93
3.13	Resultados do modelo de regressão logística com efeitos aleatórios em [-or] e [-ar] para grupo de análise.	94
3.14	Resultados parciais do modelo de regressão logística com efeitos aleatórios em [-or] e [-ar] para pseudopalavras.	96
3.15	Atribuições acentuais por <i>pseudopalavra</i> em [-or] e [-ar] (em ocorrências).	97
3.16	Atribuições acentuais por <i>repetição</i> em [-or] e [-ar].	98
3.17	Atribuições acentuais proparoxítonas por <i>grupo de análise</i>	102
3.18	Atribuições acentuais proparoxítonas por <i>pseudopalavra</i> (em ocorrências).	103
3.19	Atribuições acentuais por <i>informante</i>	104
3.20	Atribuições acentuais do grupo iii por <i>grupo de análise</i>	105
3.21	Informações sociais de uma amostra de informantes.	106
1	Resultados gerais de atribuição acentual por <i>condição experimental</i> e <i>pseudopalavra</i>	125
2	Pseudopalavras por sílaba/segmento alvo.	128
3	Atribuição acentual no pré-teste.	128
4	Atribuição acentual por pseudopalavra no pré-teste.	129

Lista de Quadros

1.1	Predições de atribuição acentual por hipótese de análise. . . .	24
2.1	Hipóteses do Grupo de Análise 1.	57
2.2	Hipóteses do grupo de análise 2.	60
2.3	Variáveis dependente e independentes com suas propriedades e variantes do Grupo de Análise 1.	66
2.4	Variáveis dependente e independentes com suas propriedades e variantes do Grupo de Análise 2.	67
3.1	Predições de atribuição acentual por hipótese de análise. . . .	70
3.2	Hipóteses dos grupos de análise 1.	72
3.3	Hipóteses do Grupo de Análise 2.	88

Lista de Figuras

1.1	Espectograma de duração e intensidade de vogal tônica em [ka.mi'tor].	6
1.2	Representação de redes e esquemas acentuais no latim (CAN- TONI, 2013, p. 67).	21
1.3	Mapa de Uso (CANEVER, 2012, p. 14).	27
2.1	Exemplo de imagens prototípicas empregadas para incitar a pseudopalavra [si.ta.ba].	42
2.2	Exemplo do experimento com a pseudopalavra [ba.zi.gar].	63
3.1	Esquema de conexões fonológicas entre [niw] e [tʃiw].	76
3.2	Adequação do modelo de regressão logística em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw].	79
3.3	Erros do modelo de regressão logística binomial com efeitos mistos em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw].	80
3.4	Esquema de conexões fonológicas entre [-or] e [-ar].	91
3.5	Atribuição acentual por sílaba final em [-or] e [-ar].	96
1	Resultados do modelo de regressão logística binomial do meta- experimento.	126
2	Gráfico de adequação do modelo de regressão logística no Grupo de Análise 1	130
3	Erros do modelo de regressão logística binomial com efeitos mistos em [niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw].	131

4	Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 1 no modelo de [grupo de análise + sílaba final + repetição].	132
5	Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 1 no modelo de [pseudopalavra + repetição].	133
6	Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 2 no modelo de [grupo de análise + sílaba final + repetição].	134
7	Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 2 no modelo de [pseudopalavra + repetição].	135

Introdução

Esta dissertação tem por objetivo investigar o papel que a frequência dos tipos silábicos desempenha na atribuição acentual em pseudopalavras por falantes nativos do Português Brasileiro (doravante PB). Para atingir tal propósito, avaliaram-se, por meios experimentais, as tendências para a atribuição do acento primário em pseudopalavras a partir de tipos silábicos específicos. A principal contribuição desta dissertação é lançar luz ao debate sobre a atribuição do acento lexical no PB através da avaliação de quatro propostas sugeridas pela literatura: Hipótese Lexical (CÂMARA JR., 1970 [2001]), Hipótese Métrica (BISOL, 1994), Hipótese Morfológica (LEE, 1995) e Hipótese Multirrepresentacional (CANTONI, 2013). Argumenta-se que a perspectiva teórica multirrepresentacional, a qual engloba os Modelos Dinâmicos baseados no Uso (KEMMER & BARLOW, 1999; LANGACKER, 1999), a Fonologia de Uso (BYBEE, 2001; 2006; 2010), a Teoria de Exemplares (BYBEE, 2010; PIERREHUMBERT, 2001) e a Linguística Cognitiva (LANGACKER, 1999; 2008; 2013), oferece contribuições relevantes para a compreensão do acento lexical no PB.

Vale ressaltar que a presente pesquisa restringe o seu objeto de análise ao acento em não verbos, posto que o comportamento acentual nos verbos é regular, os quais seguem o padrão acentual do tempo e do modo verbal, independentemente de sua conjugação (1ª, 2ª ou 3ª). Deve-se a tal fato que as propostas de análise apresentadas nesta dissertação limitam-se aos não verbos.

Com a finalidade de avaliar e quantificar os tipos silábicos do PB, um *corpus* linguístico foi compilado, o qual foi denominado de *Corpus* ABG¹. A

¹ O *Corpus* ABG resulta de um trabalho em parceria com o colega Bruno Ferrari Guide.

metodologia empregada na seleção dos textos, extração e quantificação dos dados pautou-se nos pressupostos da Linguística de *Corpus* (cf. BIBER, 1993; MCENERY & WILSON, 1996; SARDINHA, 2000a; 2000b). A compilação do *corpus* permitiu a extração da frequência de tipo e de ocorrência de diversos padrões fonológicos, como da palavra, da sílaba, dos segmentos e das estruturas silábicas. Para que essas informações pudessem ser extraídas, desenvolveram-se ferramentas computacionais (silabificador, transcritor, acentuador) em Linguagem Python (cf. Python Team, 2015).

A partir do exame do *Corpus* ABG, foram constituídas pseudopalavras a serem submetidas às condições experimentais. Segundo Capovilla, Varanda & Capovilla (2006), pseudopalavras podem ser compreendidas como “seqüências de caracteres que compõem um todo pronunciável, mas carente de significado” (p. 48), as quais devem respeitar todas as restrições fonotáticas da língua. Em princípio, parece ser contraditório investigar a atribuição do acento lexical em pseudopalavras, visto que, como as palavras não existem, a atribuição do acento poderia ser ao acaso. Contudo, é essa aparente contradição que motiva o presente trabalho. É exatamente porque os falantes têm gramática fonológica que é possível avaliar os princípios de atribuição do acento lexical em pseudopalavras. De fato, a partir de sua gramática fonológica os falantes podem atribuir o acento lexical a qualquer palavra que lhes seja apresentada. Basicamente, a presente dissertação apresenta os seguintes questionamentos: Qual é o comportamento acentual dos falantes diante de vocábulos novos? Os correlatos fonéticos acentuais utilizados em pseudopalavras são os mesmos de palavras reais? Quais informações os indivíduos avaliam em sua gramática fonológica a fim de atribuir tonicidade a um vocábulo: o peso silábico, a ausência de um marcador de palavra, a frequência dos padrões fonológicos e/ou os processos analógicos?

Para que a pesquisa pudesse se realizar, selecionaram-se dois grupos experimentais: i. [niw], [V_{oral}.tʃiw], [V_{nasal}.tʃiw]; e ii. [-or], [-ar]. O grupo de análise 1, composto pelos tipos silábicos [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], investigou, por meio do conceito de lacuna fonológica², como a frequência

² A lacuna fonológica, como será denominada neste trabalho, diz respeito à ausência ocasional de uma seqüência sonora, com uma certa tonicidade, na gramática fonológica do falante. Para mais detalhes,

dos tipos silábicos impacta na determinação da localização acentual. Mais especificamente, verificou-se se sequências sonoras que no PB possuem somente tipos silábicos oxítonos, como [niw] (i. e., *anil*, *canil*) e [V_{nasal}.tʃiw] (i. e., *cantil*, *infantil*), conduzem a produções acentuais exclusivamente oxítonas em pseudopalavras. Em contrapartida, se sequências sonoras que possuem tanto acento paroxítono como oxítono, como [V_{oral}.tʃiw] (i. e., *versátil*, *sutil*), apresentariam produções paroxítonas e oxítonas.

O grupo de análise 2, composto pelos tipos silábicos [-or] e [-ar], verificou em que medida a frequência dos tipos silábicos interage com processos analógicos, tendo como motivação a frequência de vocábulos fonológicos similares. Mais especificamente, investigou-se se a sequência sonora [-or] receberia um maior número de atribuições acentuais paroxítonas do que a sequência [-ar]. Era esperado um maior número de produções paroxítonas em [-or] em decorrência de ele apresentar vocábulos fonológicos similares frequentes com esse padrão acentual - apesar de ser um tipo infrequente com tonicidade paroxítona. Em contrapartida, em [-ar], por ele ser altamente produtivo com acento oxítono e possuir poucos e infrequentes vocábulos fonológicos paroxítonos, pressupôs-se preferência pelo padrão acentual oxítono.

Esta dissertação estrutura-se da seguinte maneira: o capítulo 1 dedica-se à exposição do percurso teórico que fundamenta esta pesquisa. Para tal, apresentam-se o acento numa perspectiva fonética e fonológica, a relação entre sílaba e acento e, ainda, os pressupostos que norteiam os Modelos Multirrepresentacionais.

O capítulo 2 apresenta os princípios metodológicos que pautaram a seleção da condição experimental e a compilação de um *corpus* linguístico, bem como as motivações que conduziram à adoção dos tipos silábicos [niw], [V_{oral}.tʃiw], [V_{nasal}.tʃiw], [-or] e [-ar]. Expõem-se, também, o perfil dos informantes, o método de formação dos estímulos e o método de análise dos dados, os quais consistem em outras decisões metodológicas relevantes para a realização do desenho experimental.

O capítulo 3 discute os resultados obtidos nos grupos de análise 1 e 2, bem como o estatuto fonológico das proparoxítonas, as quais, como será visto,

veja a seção 2.2.2, do capítulo 2.

emergiram nos dois grupos de análise, e o papel dos indivíduos nos estudos acentuais. Por fim, no capítulo 4, explicitam-se as conclusões alcançadas a partir dos experimentos realizados e as contribuições que a presente pesquisa fornece aos estudos acentuais no PB. Seguem-se a bibliografia e os anexos.

1

Pressupostos Teóricos

Este capítulo reserva-se à descrição dos pressupostos teóricos da presente pesquisa. O capítulo está estruturado da seguinte maneira: na seção 1.1, apresentam-se os conceitos para a compreensão do acento primário em português na perspectiva fonética; na seção 1.2, explicitam-se duas propostas de representação das sílabas e suas implicações para a atribuição do acento primário; na seção 1.3, sintetizam-se as principais propostas de análise para o acento primário em português numa perspectiva fonológica, expondo os méritos e as lacunas de cada uma delas; por fim, na seção 1.4, exploram-se os pressupostos dos Modelos Multirrepresentacionais, os quais englobam os Modelos Dinâmicos baseados no Uso, a Fonologia de Uso e a Teoria dos Exemplares.

1.1 O acento na perspectiva fonética

Estudos dedicados à investigação da natureza fonética do acento primário em Português Brasileiro apontam os seguintes correlatos físicos para o acento: duração, intensidade e frequência (cf. MARTINS, 1988; MASSINI-CAGLIARI, 1992; CONSONI, 2010).

Para Massini-Cagliari (1992, p. 18), "em 90% dos casos, a sílaba tônica é caracterizada por sua maior duração, o que é suficiente para fazer crer que a duração deva ser o principal correlato físico do acento em português do Brasil". Segundo Ladefoged & Johnson (2011), mesmo numa análise

perceptiva, os ouvintes percebem as sílabas tônicas com maior facilidade, já que elas frequentemente têm a vogal mais longa do que a sílaba átona, contudo, isso não significa que todas as vogais longas são, necessariamente, acentuadas.

As vogais tônicas, além de apresentar uma maior duração, são produzidas com uma maior intensidade, ou energia de emissão, o que as torna mais salientes em relação às demais vogais. Para Mateus, Falé & Freitas (2005, p. 140), "a duração refere-se ao tempo de articulação de um som, de uma sílaba ou de um enunciado". As vogais tônicas caracterizam-se por possuir uma duração intrínseca, isto é, cada segmento apresenta uma duração média. Já a intensidade "refere-se à quantidade de energia transportada pela onda de uma área percorrida num determinado período de tempo" (MATEUS, FALÉ & FREITAS, 2005, p. 102).

Como demonstra a Figura 1.1, em [ka.mi.tor], [tor] caracteriza-se como a sílaba tônica, uma vez que a vogal [o] apresenta tanto maior intensidade quanto maior duração do que as demais vogais da pseudopalavra.

Figura 1.1 Espectograma de duração e intensidade de vogal tônica em [ka.mi'tor].

A proeminência acentual, que conjuga duração e intensidade, consiste nos parâmetros acústicos a serem observados na análise dos dados deste estudo, visto que é esperado que os falantes utilizem as mesmas propriedades físicas

empregadas nas palavras reais para marcar a sílaba tônica das pseudopalavras. Ou seja, as sílabas tônicas das pseudopalavras deverão ser mais longas e ter maior intensidade do que as sílabas átonas. Semelhante comportamento fonético é presumido já que as pseudopalavras, por obedecerem todas as restrições fonotáticas do PB, poderiam ser vocábulos reais, apenas não são por ausência de significado.

1.2 Sílaba & Acento

Esta seção dedica-se à apresentação dos conceitos necessários para a compreensão da relevância da sílaba nos estudos acentuais. Resenha-se, para isso, a proposta de sílaba de Selkirk (1982), contrapondo-a com a proposta de Stuart (1982), o qual permite formular predições para a atribuição do acento em pseudopalavras.

Para Selkirk (1982), a sílaba se organiza de modo hierárquico, como apresentado a seguir:

(SELKIRK, 1982, p. 341)

Segundo a estrutura (1), a sílaba é composta de Ataque (A) e Rima (R), a qual se ramifica em Pico ou Núcleo (Nu) e Coda (Co). Todos esses elementos, com exceção do núcleo, podem ser vazios. A organização hierárquica prevê relações mais estreitas entre os elementos componentes da rima, isto é, núcleo e coda, do que entre núcleo e ataque. A relação tem como evidência alguns processos fonológicos que atuam apenas no nível da rima, como a sensibilidade ao peso silábico em alguns sistemas acentuais.

Quando a métrica de uma língua é sensível ao peso silábico, as sílabas pesadas receberão sistematicamente o acento tônico. São chamadas de sílabas

pesadas (ou ramificadas) as sílabas que possuem mais de um elemento na rima, como em (2) e (3). A sílaba é considerada ramificada quando é composta por [vogal + glide] ou por [vogal + consoante]. Para Selkirk (1982), a composição do ataque, simples ou ramificado, é irrelevante para a caracterização de sensibilidade ao peso silábico.

- (2) *Sílaba pesada por ramificação da rima*

- (3) *Sílaba pesada por ramificação do núcleo*

Em contrapartida, as sílabas que possuem um único elemento na rima são denominadas de sílabas leves, ou não ramificadas, como em (4).

- (4) *Sílaba leve*

O tipo de ramificação da estrutura silábica, se do núcleo ou da rima, consiste em outro fator relevante para a determinação de sensibilidade ao peso silábico nos sistemas acentuais. Algumas línguas, como o Mbyá (cf. MARTINS, 2006), apresentam sensibilidade à ramificação do núcleo (CVV), enquanto outras línguas, como o latim (cf. TEYSSIER, 2007), são sensíveis à ramificação

da rima (CVV ou CVC). A semelhança no comportamento de CVV e CVC, no português, em relação ao peso silábico, fez com que a presente pesquisa neutralizasse a distinção entre essas estruturas silábicas.

Contrapondo-se à proposta de Selkirk (1982), Stuart (1982) argumenta em prol de uma estrutura silábica flutuante, em que os elementos constitutivos, isto é, ataque, núcleo e coda, não seriam organizados de modo hierárquico universalmente, mas eles atuariam como uma unidade, com os seus elementos interagindo - ataque e núcleo, núcleo e coda, ataque e coda.

(5) *Sílaba flutuante*

(adaptação de STUART, 1982, p. 3)

A proposta de Stuart (1982) baseia-se em línguas em que a presença ou a ausência do ataque silábico determina a localização acentual, como ocorre em Aranda e Yidin (línguas australianas). Além disso, o autor demonstra que a coda não é o único elemento relevante para a determinação da duração do núcleo, posto que estudos do francês, do inglês e do coreano evidenciam que o ataque da sílaba seguinte também contribui para a determinação da duração vocálica em encontros consonantais heterossilábicos (V\$C).

Os argumentos apresentados por Stuart (1982) trazem questionamentos teóricos importantes para a proposta de Selkirk (1982), na medida em que a noção de rima como um constituinte universal e localizado numa posição privilegiada da sílaba é inquirida. A presunção de universalidade da rima e de sistemas acentuais sensíveis ao peso silábico desconsidera qualquer influência do ataque silábico, ou mesmo de informações segmentais, fato contrário aos dados encontrados por Stuart (1982).

Diante dessas considerações, a presente pesquisa assume como um dos seus pontos de investigação a relevância do peso silábico para a atribuição acentual em pseudopalavras. Para que o estudo pudesse ser realizado, selecionou-se a estrutura de palavra CV-CV-CVC. A predição é de que, se o português

for realmente sensível ao peso silábico, isto é, se considerar a ramificação da rima para a localização do acento, as produções acentuais de pseudopalavras serão categoricamente oxítonas, posto que a sílaba final ramificada atrairia o acento¹. Em contrapartida, se o acento em português não for sensível ao peso silábico, as produções acentuais de pseudopalavras serão variáveis, e não apenas oxítonas. Sugere-se que, neste caso, a atribuição do acento segue tendências dos padrões fonológicos da língua portuguesa e este é o tema desta dissertação. Prediz-se, assim, que os falantes se valeriam dos padrões abstratos presentes em suas representações mentais a fim de atribuir o acento em pseudopalavras. Vale salientar que os padrões fonológicos investigados correspondem às sílabas (ou tipos silábicos).

1.3 O acento na perspectiva fonológica

Esta seção reserva-se à apresentação e à discussão de quatro hipóteses de análise para o acento primário em português a partir de uma perspectiva fonológica. Dentre as inúmeras análises que objetivam explicar o acento primário em português (cf. CÂMARA JR, 1970 [2001]; BISOL, 1994; LEE, 1995, 2007; MASSINI-CAGLIARI, 1999; MATEUS & D'ANDRADE, 2002; CANTONI, 2013; entre outros), por motivos de concisão, foram selecionadas três propostas de análise que Ferreira-Netto (2007) considera as principais para o acento em PB, as quais consistem em: Lexical (CÂMARA JR., 1970 [2001]), Métrica (BISOL, 1994) e Morfológica (LEE, 1995), e, adicionalmente, foi considerada a proposta baseada nos Modelos Multirrepresentacionais (CANTONI, 2013). As propostas serão apresentadas a seguir.

1.3.1 Hipótese Lexical (CÂMARA JR., 1970 [2001])

A hipótese Lexical fundamenta-se nos estudos de Câmara Jr. (1970 [2001]), para o qual o acento consiste em "uma maior força expiratória, ou intensidade de emissão, da vogal de uma sílaba em contraste com as demais vogais silábicas" (p. 63), podendo incidir em qualquer uma das três sílabas a partir

¹ Mais discussões a respeito desta perspectiva de análise serão apresentadas na seção 1.3.2.

da margem direita da palavra, como em [ˈsa.bi.ɐ], [saˈbi.ɐ] e [sa.biˈa], ou, mais raramente, na quarta sílaba, como em [ˈtɛ.ki.ni.ku].

Para o autor, o acento em português tem como função ser delimitativo, distintivo e livre. Seria delimitativo já que a sua presença marca o vocábulo fonológico. O assinalamento ocorreria por meio de uma pauta acentual (veja (6)), a qual possuiria três graus. A maior intensidade, 3, designaria as sílabas mais proeminentes, as tônicas, como em (7). Casos de junção fonológica, como em (8), em que a sílaba tônica do primeiro vocábulo tem maior intensidade do que as demais sílabas pré-tônicas e, no entanto, menor intensidade do que a sílaba tônica do último vocábulo, receberiam grau 2². As sílabas pré-tônicas e pós-tônicas possuiriam 1 e 0 graus, respectivamente, já que as pós-tônicas têm menos força do que as pré-tônicas. A delimitação de vocábulos seria determinada pela presença de um 0 seguido de 1 ou de 3 seguida de 2.

(6) Pauta Acentual:

... (1) + 3 + (0) + (0) + (0)

(CÂMARA JR., 1970 [2001], p. 63)

(7) /abilidade/ *habilidade*

1 11 3 0

(CÂMARA JR., 1970 [2001], p. 63)

(8) /abilidade/ *hábil idade*

2 0 1 3 0

(CÂMARA JR., 1970 [2001], p. 63)

A mudança de significado motivada pela alteração da posição acentual na palavra confere ao acento a função distintiva, como em *jaca/jacá*, *cáqui/caqui*, a qual pode resultar em uma mudança da classe morfológica, como em “*rotulo*: *rotulo* (*verbo rotular*), *fábrica*: *fabrica* (*verbo fabricar*), *intérprete*: *interprete*

² Para Câmara Jr. (1970 [2001], p. 63), “numa seqüência de vocábulos sem pausa (o que de acordo com Paul Passy se pode chamar um grupo de força) as sílabas tônicas que precedem o último vocábulo baixam a uma intensidade 2 (...)”.

(*verbo interpretar*)” (CÂMARA JR., 1970 [2001], p. 65).

O caráter livre do acento decorreria da ausência de terminações fonológicas que definem as posições em que o acento incide. Para o autor, certas terminações podem ser mais frequentes do que outras, como os oxítonos terminados em *-r*, *mulher*, *colher*, *cobertor*, *vendedor*, mas haverá casos como *açúcar* e *revólver*, em que a entrada tardia no léxico da língua teria determinado novos padrões acentuais. Apesar do acento ser livre, a recorrência do padrão acentual paroxítono imprime à língua seu ritmo grave, isto é, trocaico.

A hipótese Lexical ressalta propriedades fundamentais que caracterizam o acento primário em PB, a saber, o caráter delimitativo e distintivo. Entretanto, por mais que Câmara Jr. (1970 [2001]) utilize o valor demarcativo do acento para distinguir os verbos dos não verbos, ele não confere aos vocábulos um tratamento diferenciado conforme a classe gramatical. Dessa forma, a regularidade na atribuição do acento verbal e a imprevisibilidade do acento nominal não são consideradas pelo autor.

A proposta lexical, ainda, ao propor que o acento seja livre, assume que o acento pode incidir em qualquer posição da palavra desde que obedeça a janela de três sílabas a partir da margem direita da palavra, isto é, para esta proposta, o acento seria atribuído ao acaso, não acolhendo as tendências fonológicas, morfológicas e semânticas da língua. Do ponto de vista fonológico, há uma clara preferência do acento por palavras terminadas em consoantes ou ditongos, como em *colher* e *troféu*. No âmbito da morfologia, a produtividade das terminações verbais *-ar*, *-er* e *-ir*, por exemplo, podem direcionar os falantes a interpretarem vocábulos terminados em *-r* como verbos. Já do ponto de vista semântico, algumas terminações ou sequências fonológicas contêm traços que remetem a certos campos semânticos. Por exemplo, a recorrência de palavras terminadas em [ɔw] relacionadas a nomes de remédios, *paracetamol*, *diclofenol*, *claridol*, faria com que o falante, ao escutar uma palavra como [sistikari'mɔw], a associe ao campo semântico de remédios.

A análise de Câmara Jr.(1970 [2001]) tem o mérito de ser a primeira análise de cunho linguístico sobre o acento do português, a qual buscou elencar argumentos para a atribuição lexical do acento: cada palavra teria o seu padrão acentual, o qual poderia ser distintivo, ou fonêmico em relação a

outros padrões acentuais. No contexto histórico em que se insere a proposta do autor, o caráter contrastivo do acento era o ponto central, de maneira análoga ao contraste observado em pares mínimos dos sons. Enquanto, de um lado, os pares *sábia*, *sabia* e *sabiá* elucidavam o contraste acentual entre proparoxítonas, paroxítonas e oxítonas; do outro lado, pares mínimos como *bala* e *bela* evidenciavam os fonemas [a] e [ɛ]. A proposta de Câmara Jr. (1970 [2001]) pautada na hipótese Lexical apresenta limitações dentre as quais se destacam: i. não considerar a regularidade na atribuição do acento verbal em relação a não previsibilidade do acento nominal; e ii. não acolher as tendências segmentais e semânticas da língua.

1.3.2 Hipótese Métrica (BISOL, 1994)

A hipótese Métrica consiste em um avanço teórico fundamentado na Fonologia Métrica, a qual foi instituída por Liberman (1975) e desenvolvida por Halle & Vergnaud (1987), como parte da fonologia não linear. Um dos pontos centrais da fonologia não linear consistiu em apresentar propostas que explicassem a organização segmental das sílabas a partir dos constituintes silábicos: Ataque (A), Rima (R), Núcleo (Nu) e Coda (C).

De maneira análoga à proposta de sílaba de Selkirk (1982), apresentada na seção 1.2, a Fonologia Métrica considera que os constituintes simples, ou sílabas leves, possuem uma única posição esquelética projetada na rima, enquanto os constituintes compostos, ou sílabas pesadas ou ramificadas, possuem ao menos duas posições esqueléticas na rima. As posições esqueléticas são representadas pela projeção de (x), como em (9).

Considere a representação da palavra *bailar* em (9), a qual apresenta duas sílabas. Tanto a primeira sílaba com núcleo ramificado como a última sílaba com rima ramificada são consideradas sílabas pesadas, as quais, em algumas línguas, desempenham um papel na atribuição do acento lexical. Vale ressaltar que embora todos os segmentos projetem (x), para a Fonologia Métrica, somente a projeção de duas posições esqueléticas dentro da rima caracteriza uma sílaba pesada em línguas que assumem sensibilidade ao peso silábico.

A Fonologia Métrica oferece instrumentos para explicar a interação entre a estrutura silábica e a atribuição do acento a partir da alternância de sílabas fortes (*s = strong*) e sílabas fracas (*w = weak*), a qual se configura em pés métricos (*s w*) ou (*w s*). A posição em que a sílaba forte se encontra demarca a cabeça de proeminência do pé métrico, o qual é denominado de troqueu se apresentar proeminência à esquerda (*s w*) ou de iambo se apresentar proeminência à direita (*w s*).

Segundo Bisol (1994), o constituinte prosódico que gera o algoritmo acentual no PB é o pé métrico binário troqueu (*s w*) ou (*.), em que o asterisco indica a cabeça do pé métrico. Assim, na palavra ['ka.zɐ], o pé troqueu (*s w*) é formado pela sílaba forte ['ka] seguida pela sílaba fraca [zɐ]. Para a autora, duas regras são suficientes para gerar o acento, conforme (10), a ser aplicada na palavra derivacional para os não verbos:

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

(BISOL, 1994, p. 25)

A regra (i) prediz que as sílabas finais pesadas atraem o acento, fato que explica a tonicidade oxítona em vocábulos como *amor* e *troféu*. A regra (ii), por sua vez, indica que, caso não haja sílabas pesadas, o padrão trocaico é favorecido. Explica-se, assim, o padrão acentual paroxítono em palavras como

parede e borboleta.

As regras de (10) não conseguem explicar todos os padrões acentuais do PB, visto que vocábulos como *médico* e *versátil* não seguem os padrões acentuais propostos pela autora. Faz-se necessária, dessa forma, a postulação de um mecanismo denominado de extrametricidade, o qual, segundo Collischonn (2010, p. 151), refere-se a um “recuo do acento uma sílaba [ou segmento] à direita da sua posição esperada”. Esse recurso tem como função o ajuste da palavra prosódica às regras gerais por meio da utilização do <>, o qual indica as sílabas ou os segmentos que são invisíveis às regras acentuais, isto é, que são extramétricos. A extrametricidade atinge, portanto, "i. palavras com acento na terceira sílaba; [e] ii. palavras terminadas em consoante ou ditongo com acento não-final" (BISOL, 1994, p. 27), como visualizado em (11) e (12).

(11) *Extrametricidade silábica:*

fósforo	árvore	número
/fɔsfor+o/	/arvor+e/	/numer+o/
<ro>	<re>	<ro>
(* .)	(* .)	(* .)

(adaptação de BISOL, 1994, p. 28)

(12) *Extrametricidade de coda silábica:*

útil	fácil	visível
/util/	/fasil/	/vizivel/
<l>	<l>	<l>
(* .)	(* .)	(* .)

(adaptação de BISOL, 1994, p. 29)

O principal problema encontrado na postulação de mecanismos de extrametricidade consiste em que não há como determinar de antemão, por exemplo, o porquê de *hípe*<*r*> conter um segmento extramétrico e receber acento paroxítono, ao passo que, em *amor*, o segmento *r*, que também está em posição final, não é considerado extramétrico e recebe acento oxítono. Note que a postulação de segmentos ou sílabas invisíveis às regras acentuais requer

marcação lexical que estabeleça quais vocábulos são submetidos ou não à extrametricidade. Ou seja, uma estipulação ao invés de uma explicação.

Segundo Bisol (1994), o maior problema para a proposta métrica consiste nas palavras acentuadas na sílaba final que não possuem rima ramificada, isto é, as oxítonas com sílaba final leve, como em *café*, *caju*, *jacaré*. Para as palavras desta categoria, a análise busca assemelhar-se às palavras com acento e rima ramificada em posição final, por meio da postulação de uma consoante final subjacente que só se manifesta nas formas derivadas, como em (13):

(13) *Consoante Final Abstrata*

Abecê - abece**D**ário

Café - cafe**T**eira

Chá - cha**L**eira

Pé - pe**D**al

Só - so**L**idão

(BISOL, 1994, p. 30)

Dentre os méritos da hipótese Métrica, têm-se a justificativa, por meio de apenas dois algoritmos, dos dois padrões acentuais mais recorrentes da língua, à medida que se reconhece também o papel de sequências segmentais, ditongos e consoantes pós-vocálicas, e do pé métrico trocaico na atribuição acentual do PB. Contudo, as principais limitações consistem em: i. estipular de maneira *ad hoc* a atribuição de extrametricidade e a emergência da consoante subjacente; e ii. desconsiderar que há subpadrões no sistema acentual do PB, os quais não podem ser expressos por meio de regras determinísticas.

1.3.3 Hipótese Morfológica (LEE, 1995)

A hipótese Morfológica apresentada nesta seção baseia-se nos estudos de Lee (1995), para o qual o acento primário no PB não seria sensível ao peso silábico, mas sim à categoria lexical da palavra. Para o autor, verbos e não verbos possuem comportamentos acentuais diferentes. Dessa maneira, domínios

distintos seriam requeridos para a aplicação das regras acentuais: no nível α , aplica-se a regra dos não verbos, e, no nível β , dos verbos.

Baseado na proposta de Mateus (1983 apud LEE, 1995), Lee propõe que o acento em não verbos incide na última vogal do radical, como em (14). Para isso, postula uma extrametricidade morfológica dos marcadores de palavra (vogal temática), que ocorreria no nível α , impedindo sua visibilidade nos domínios prosódicos.

- (14) A Estrutura do Não-Verbo
 Radical + {Vogal Temática}

(LEE, 1995, p. 152)

A estrutura dos não verbos permite que vocábulos como as paroxítonas com sílabas finais leves (15-a), as oxítonas com sílabas finais pesadas (15-b) e as oxítonas com sílabas finais leves (15-c) sejam explicados.

- (15) a. [almoç]o, [menin]o
 b. [tonel], [amor]
 c. [café], [caqui]

A regra formulada para esses casos pode ser verificada em (16):

- (16) Regra de Acento do Não-Verbo (não marcado)
 Domínio: radical derivacional
 a. Constituintes binários
 b. Não-iterativo
 c. Cabeça à direita
 d. Direita para Esquerda

(LEE, 1995, p. 153)

Fundamentado no poder explicativo que o algoritmo (16) tem em relação às paroxítonas e às oxítonas, Lee confere a elas a concepção de padrões acentuais não marcados. O conceito de marcação fundamenta-se nas propriedades

linguísticas dos fenômenos sonoros, em que "uma propriedade pode ser dita não marcada se é recorrente nas línguas naturais e é dita marcada se tem caráter excepcional e ocorrência marginal nas línguas naturais" (CRISTÓFARO-SILVA, 2011, p. 149). Os acentos paroxítono e oxítono seriam não marcados, visto que são mais comuns no sistema acentual do PB, ou seja, apresentam maiores índices de frequência. Em contrapartida, casos como as paroxítonas com sílaba final pesada, *túnel* e *jovem*, e as proparoxítonas, *último* e *árvore*, que são menos frequentes na língua, são denominados de padrões acentuais marcados.

O critério de marcação consiste no mecanismo adotado por Lee (1995) para diferenciar os vocábulos que são incluídos ou não na regra (16). Com isso, aos padrões acentuais em que (16) encontra problemas, já que o acento não incide na última vogal do radical, outra regra tem que ser postulada, a qual se destina aos não verbos marcados, como expõe (17).

(17) Regra de Acento do Não-Verbo (marcado)

Domínio: radical derivacional

- a. Constituinte binário
- b. Cabeça à esquerda
- c. Direita para esquerda
- d. Não-iterativo

(LEE, 1995, p. 154)

De modo semelhante à extrametricidade proposta por Bisol (1994), a postulação de regras destinadas a vocábulos marcados ou não marcados inviabiliza que se saiba de antemão o porquê de *híper* ser submetido à regra de palavras marcadas e *amor*, à regra de não marcadas.

Dentre os méritos da hipótese Morfológica, tem-se como contribuição principal a utilização de informações linguísticas que não se limitam às fonológicas. Conferir às palavras de classes morfológicas diferentes um tratamento diferenciado conforme a sua categoria acentual consiste em uma justificativa plausível diante da diversidade de comportamento entre verbos e não verbos. No entanto, a proposta de Lee (1995) traz a problemática de distinguir os

não verbos entre marcados e não marcados. Do ponto de vista morfológico, não há razões para a diferenciação dos não verbos, posto que, por exemplo, tanto *abóbora* e *jovem*, vocábulos marcados, quanto *café* e *almoço*, vocábulos não marcados, compõem uma mesma classe gramatical, a de substantivos.

Apesar de a hipótese Morfológica tecer considerações relevantes no que se refere à distinção do comportamento entre verbos e não verbos, as principais limitações encontradas na proposta consistem em: i. estipular de maneira *ad hoc* a atribuição de marcação; e ii. diferenciar vocábulos marcados e não marcados sem que haja implicações morfológicas para tal distinção.

1.3.4 Hipótese Multirrepresentacional (CANTONI, 2013)

A hipótese Multirrepresentacional apresentada em Cantoni (2013) fundamenta-se nos pressupostos da Linguística Cognitiva (LANGACKER, 1987), dos Modelos baseados no Uso (BARLOW & KEMMER, 2000; BYBEE, 2001; 2010) e dos Sistemas Dinâmicos (BECKNER et al., 2009; BYBEE, 2010), para os quais a língua é "o 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" (CANTONI, 2013, p. 14).

A proposta de Cantoni (2013) parte de uma perspectiva diacrônica da língua a fim de explicar os padrões acentuais existentes no português. Para a autora, o tempo, em consonância com os sistemas dinâmicos, consiste em um fator determinante para o estado atual de um sistema. Dessa forma, o estado atual de uma língua é resultado de forças e atratores³ que atuaram no sistema desde o seu estado inicial. De acordo com esse ponto de vista, sendo o português proveniente da língua latina, embora seja um sistema específico, o seu sistema acentual traria resquícios do sistema acentual latino, na medida em que o uso teria construído representações abstratas que mapearam as experiências linguísticas.

O sistema acentual latino, que era regular, caracterizava-se por ser um sistema sensível ao peso silábico. O acento incidia na penúltima sílaba se

³ Atratores são estados que ocorrem com frequência, veja mais detalhes na seção 1.4.

esta fosse longa, como em *amīcum*, *capīllum*, do contrário, o acento incidia na antepenúltima sílaba, como em *árbōrem*, *hómīnem*, exceto nos dissílabos, que detinham o acento na penúltima sílaba independentemente da duração da vogal, como em *séptem*, *dátum*⁴. A interação entre peso silábico e acento era tão robusta que mesmo os vocábulos provenientes de outras línguas eram adaptados ao sistema acentual latino.

Para Cantoni (2013), por o latim apresentar um sistema acentual previsível, obedecendo a janela de três sílabas a partir da margem direita da palavra e ser sensível ao peso silábico, a regularidade e a produtividade dos padrões acentuais teriam provocado, em conjunção com as experiências linguísticas, a emergência da estrutura gramatical por meio de esquemas abstratos, conforme exposto na Figura 1.2. Isto é, a produtividade do sistema acentual latino teria provocado a fortificação das redes morfológicas.

⁴ Exemplos extraídos de Teyssier (2007, p. 9).

Figura 1.2 Representação de redes e esquemas acentuais no latim (CANTONI, 2013, p. 67).

A Figura 1.2 expõe os *templates* acentuais provenientes, justamente, da relação entre abstração, experiência de uso e estrutura gramatical. Note que, no latim, as estruturas abstraídas refletem padrões acentuais regulares e recorrentes da língua, isto é, (S x), (s S x) e (S s x), os quais são regulados pela sensibilidade ao peso (quantidade vocálica) e pela fortificação das estruturas morfológicas.

O acento no português, segundo Cantoni (2013), emerge de relações semânticas, fonéticas e morfológicas estabelecidas já no latim. A sensibilidade ao peso silábico, marcada através da quantidade vocálica, que regulava o

sistema acentual latino, a qual é perdida na passagem do latim para o português, faz com que haja uma reestruturação do sistema, provocando assim a emergência das conexões morfológicas como atratores no português. Com isso, o léxico do português, que provém, majoritariamente, de uma herança latina, teria um acento *default* em nomes não derivados e em nomes flexionados⁵, enquanto, em palavras provenientes de derivação nominal, o acento seria regido pela morfologia. Isso porque, como Cantoni (2013) apresenta, certos sufixos podem atrair o acento, *-ação*, *-ismo*, *repeli-lo*, *-ico*, *-imo*, ou mesmo preservar a tonicidade da forma base, no caso de flexão verbal, *canto*, *cantas*.

A proposta de Cantoni (2013) traz, ainda, duas predições importantes:

[i.] não haveria um mecanismo único responsável pela localização do acento no português. [e ii.] As aparentes relações entre estrutura silábica e localização do acento podem ser interpretadas como resquício dos *templates* acentuais do sistema latino (CANTONI, 2013, p. 72-73).

A primeira afirmativa dialoga com a perspectiva dos Modelos Multirrepresentacionais, para os quais diversos fatores são responsáveis pela construção dos esquemas abstratos presentes na mente do falante. Já a segunda afirmativa justifica a constante correlação atribuída ao acento oxítono e às palavras com sílaba final pesada (cf. BISOL, 1994).

Para os Modelos Multirrepresentacionais, as generalizações, que são realizadas a partir de exemplares de uso e provocam a emergência de esquemas, expressam padrões abstratos recorrentes na língua. A esquematização dos padrões, que compõe a gramática do falante, permite com que ele utilize e estenda o padrão experienciado a novas construções da língua. São justamente os padrões abstratos que possibilitam a produção de vocábulos desconhecidos ou mesmo de pseudopalavras. Para Cantoni, por meio de generalizações, o falante replicaria os padrões abstratos existentes na língua aos novos vocábulos.

⁵ Para a hipótese Multirrepresentacional, o morfema de plural não modificaria a localização acentual da forma base.

A hipótese Multirrepresentacional tem como um de seus méritos a possibilidade de que metodologias experimentais possam ser utilizadas⁶. Experimentos podem investigar, por exemplo, a proposta de que padrões abstratos emergem de generalizações de experiências de uso que são armazenadas e reutilizadas em novos vocábulos, o que possibilita que todos os padrões acentuais se manifestem. O presente estudo, portanto, assume a proposta de Cantoni (2013) como ponto inicial para a realização de uma pesquisa mais ampla que investiga a aplicação dos padrões acentuais em pseudopalavras. Entretanto, ressalvas devem ser feitas em relação aos Modelos Multirrepresentacional de maneira geral, visto que não é possível predizer como se dão e em quais níveis (segmental, silábico, lexical) ocorrem os processos de abstração. Resta explicar também como diferentes experiências de uso, as quais apresentam diferentes índices de frequências, possibilitam que generalizações ocorram, e, ainda, quais os gastos de memória necessários, em termos de processamento, para que se tenha uma atribuição de acento lexical, mesmo que as atribuições ocorram a partir de processos de abstração em redes.

Síntese

O Quadro 1.1 sintetiza as predições acentuais que cada hipótese de análise apresentada ao longo deste capítulo faz para a atribuição acentual em pseudopalavras em PB.

⁶ Para as hipóteses Métrica e Morfológica, a utilização de metodologias experimentais em pseudopalavras implica a exclusão das regras de extrametricidade, de consoante final abstrata e de marcação lexical, o que, de certa forma, restringe as produções acentuais a regras categóricas - fato improvável de ocorrer.

Hipótese	Predições para atribuição de acento em pseudopalavras
Lexical	Não há predição para a atribuição de acentoônico em pseudopalavras. Isto porque de acordo com a hipótese Lexical cada palavra tem seu próprio acento, o que resultaria numa atribuição acentual aleatória, sem qualquer tendência específica.
Métrica	A hipótese Métrica prediz que sílabas pesadas tendem a receber acento. Assim, espera-se que o padrão acentual em pseudopalavras seja oxítono se houver sílaba pesada e paroxítono (troqueu) se não houver sílaba pesada.
Morfológica	A hipótese Morfológica prediz que haverá diferença na atribuição de acento entre nomes e verbos. No caso de nomes, que é o objeto de estudo desta dissertação, a predição é que o acento incida na última vogal do radical, isto é, acento paroxítono em palavras que terminam com vogal temática e oxítono nos demais casos.
Multirrepresentacional	A hipótese Multirrepresentacional prediz que abstrações refletem padrões estocásticos, inferidos a partir do uso, que refletem valores de frequência. A predição é que a atribuição acentual refletirá as tendências de frequência da língua, com isso, qualquer padrão acentual permitido no PB pode ocorrer ⁷ .

Quadro 1.1 Predições de atribuição acentual por hipótese de análise.

A partir das predições exibidas no Quadro 1.1, realizou-se um experimento com o objetivo de avaliar as hipóteses de atribuição de acentoônico em

⁷ Ao propor que o acento reflete padrões estocásticos da língua inferidos a partir do uso, a proposta Multirrepresentacional postula padrões acentuais que exprimem índices de frequência da língua, os quais não são categóricos. Diante de tal fato, seu falseamento seria possível através de resultados categóricos, seja por produções oxítonas seja por paroxítonas.

pseudopalavras. Por, até o momento, a hipótese Multirrepresentacional possibilitar uma análise mais abrangente do acento em PB, a próxima seção apresenta os pressupostos dos Modelos Multirrepresentacionais que englobam a Fonologia de Uso, a Teoria dos Exemplares, a Linguística Cognitiva e os Modelos Dinâmicos baseados no Uso, uma vez que é nesta perspectiva teórica que se enquadra esta dissertação.

1.4 Modelos Multirrepresentacionais

Os Modelos Multirrepresentacionais caracterizam-se por conceber o uso como fator primordial para a emergência da gramática dos falantes a partir de representações linguísticas múltiplas, os quais englobam os Modelos Dinâmicos baseados no Uso (KEMMER & BARLOW, 1999; LANGACKER, 1999), a Fonologia de Uso (BYBEE, 2001; 2006; 2010), a Teoria dos Exemplares (BYBEE, 2010; PIERREHUMBERT, 2001) e a Linguística Cognitiva (LANGACKER, 1999; 2008; 2013).

Esta perspectiva teórica admite que a língua é parte da cognição, emergindo por meio de instâncias de uso e de processos cognitivos gerais, compartilhados por outras capacidades psicológicas. A Gramática Cognitiva, portanto, tem como meta

descrever a natureza da linguagem como ela na verdade ocorre nos falantes, ao representar toda a complexidade e redundância que a caracteriza. O termo "cognitiva" implica que o objetivo da Gramática Cognitiva é encontrar um modelo de língua que esteja de acordo com o que é conhecido a respeito da cognição humana em geral e tentar, tanto quanto possível, excluir as entidades de análises que parecem não ter qualquer existência independente fora de outros modos de cognição (NATHAN, 2007, p. 627)⁸.

⁸ Trecho Original: "One of the major goals of Cognitive Grammar is to describe the nature of language as it actually occurs in speakers, accounting for all the complexity and redundancy that characterize it. The term "cognitive" implies that Cognitive Grammar's goal is to come up with a model of language that is in accord with what is known about human cognition in general and to attempt, as far as possible, to exclude from the analysis entities that appear to have no independent existence outside of other modes of cognition".

Segundo Langacker (1999), abstração, arraigamento, comparação, associação e composição são os fenômenos cognitivos envolvidos e essenciais à língua. A habilidade de tecer abstrações é responsável pela construção de rotinas cognitivas implementadas pelo uso real. É por meio dela que os eventos de uso são automatizados e tornam-se esquemas armazenados nas representações mentais. A cada novo uso, o esquema é reforçado, fato que facilita a sua recorrência e provoca o seu arraigamento. A força nas conexões provocadas pela repetição possibilita que as estruturas arraigadas emergjam como atratores dentro do sistema linguístico. Segundo Fleischer (2011, p. 75),

em qualquer sistema, os estados que ocorrem repetidamente ou que são aproximados com frequência e de forma cada vez mais próxima constituem um conjunto de atratores. Tais atratores são meramente estados que ocorrem com frequência, enquanto outros estados, embora plenamente possíveis, simplesmente não ocorrem.

O processo de construção dos esquemas implica a comparação e a abstração de inúmeras unidades similares (ou padrões), enquanto a esquematização constitui o estabelecimento de esquemas particulares, com diferentes níveis de granularidade (cf. LANGACKER, 1999; NATHAN, 2007). No primeiro caso, para os esquemas, as abstrações correspondem às amplas generalizações extraídas das unidades linguísticas, como no passado regular do inglês, *walk/walked*, *talk/talked*, *work/worked*. No segundo, para as esquematizações, as regularidades têm escopo limitado, ou mesmo são unidades específicas, como no passado irregular do inglês, *sing/sang*, *drink/drank*, *swim/swam*.

A elaboração dos esquemas ou esquematizações ocorre por meio de operações comparativas. Quando há correspondência, categoriza-se o padrão de comparação ao alvo, $[A] \rightarrow (B)$; quando há discrepância, extensiona-se, $[A] \dashrightarrow (B)$. De modo geral, essas ações acontecem através de associações fonológicas ou semânticas, polos da estrutura simbólica.

O último fenômeno cognitivo fundamental à língua consiste na composição, a qual, segundo Langacker (1999, p. 4, grifo do autor), "envolve a **integração** de duas ou mais estruturas de **componentes** para formar uma

estrutura **composta**⁹. A recorrência de padrões neurofisiológicos das unidades componentes conduz à fortificação e ao arraigamento dessas estruturas, solidificando-as como uma nova unidade.

Os processos cognitivos e os eventos de uso constituem o verso e o averso do conhecimento linguístico do falante. As instâncias de uso se alimentam à medida que resultam do sistema linguístico. Fator e produto, a experiência fornece subsídios para a construção de esquemas cognitivos, sendo que estes guiam o uso. Como num processo de retroalimentação, ou *loop*, gramática e uso influenciam-se e relacionam-se de modo direto, como explicitado na Figura 1.3 (cf. KEMMER & BARLOW, 1999; LANGACKER, 1999; CANEVER, 2012).

Figura 1.3 Mapa de Uso (CANEVER, 2012, p. 14).

Nessa perspectiva, a atuação dos processos cognitivos é responsável pela construção das categorias fonológicas presentes na gramática dos falantes por meio do uso. O processo de abstração faz com que as unidades fonológicas sejam automatizadas e arraigadas às representações linguísticas, conduzindo à formação de esquemas e de esquematizações, os quais refletem os padrões fonológicos presentes na língua em diferentes níveis de abstração. Os esquemas silábicos, por exemplo, emergiriam como padrões fonotáticos, ou *templates* silábicos, permitidos na língua. Eles seriam formados a partir da abstração de diversas unidades silábicas, como [liv], [hap], [mek], que conduziriam à esquematização do *template* [CVC]. De maneira análoga, padrões acentuais

⁹ Trecho original: "It involves the **integration** of two or more **component** structures to form a **composite** structure".

são abstraídos e arraigados a gramática fonológica dos falantes, como $[(\sigma \sigma)_w]$ (cf. LANGACKER, 2008; 2013). Ainda, os esquemas permitem que padrões fonológicos inexistentes emergjam por meio de processos de extensão analógica.

Na Fonologia de Uso (cf. BYBEE, 2001; 2006; 2010), os esquemas abstratos são denominados de redes ou conexões, as quais têm a frequência de uso como fator primordial para a sua construção, visto que ela determina a força das conexões e, conseqüentemente, do seu arraigamento. A frequência, segundo Bybee (2001), pode ser distinguida entre frequência de ocorrência e frequência de tipo.

A distinção empregada por Bybee pauta-se nos diferentes comportamentos que cada tipo de frequência desempenha no sistema linguístico. A frequência de ocorrência refere-se à "frequência de ocorrência de uma unidade, usualmente uma palavra" (BYBEE, 2001, p. 10)¹⁰, a qual é responsável pela preservação ou regularização de formas irregulares. Vocábulos com alta frequência de ocorrência possuem conexões fonológicas fortes, visto que o uso reforça as suas conexões, fato que conduz ao seu arraigamento e à preservação de suas estruturas, como nos paradigmas supletivos *ser*, *estar*, *ir*. Em contrapartida, os vocábulos que possuem baixa frequência de ocorrência, como os casos de participio *ganho*, *expressado*, por terem conexões fracas, são suscetíveis a nivelamentos analógicos com padrões produtivos (frequência de tipo). Além disso, a frequência de ocorrência desencadeia processos de redução fonética, a qual tende a ocorrer mais frequentemente em palavras ou sequências sonoras de alta frequência do que em sequências de média ou baixa frequência, uma vez que a repetição dos padrões neurofisiológicos faz com que as sequências sonoras sejam mais fluentes, como em $['psize]$ *precisa* e $['sêse]$ *licença*.

Quanto à frequência de tipo, ela corresponde à "frequência do dicionário de um padrão particular (como de um padrão acentual, de um afixo ou de um encontro consonantal)" (BYBEE, 2001, p. 10)¹¹. O mapeamento de diversas unidades, por meio de similaridade fonológica, faz com que esquemas ou

¹⁰ Trecho Original: "TOKEN FREQUENCY is the frequency of occurrence of a unit, usually a word".

¹¹ Trecho Original: "TYPE FREQUENCY refers to the dictionary frequency of a particular pattern (e.g., a stress pattern, an affix, or a consonant cluster)".

redes de conexões emergjam na língua, os quais refletem padrões fonológicos abstratos. Esses padrões, de maneira geral, expressam padrões regulares e produtivos da língua. As conexões (ou esquemas) se fortalecem e se arraigam às representações a partir de eventos de uso. Quanto mais forte for a conexão de um padrão, mais probabilidade ele terá de ser generalizado para padrões menos produtivos (ou menos frequentes) e de aplicar-se a novos lexemas da língua. No paradigma verbal do português, a alta produtividade da 1ª conjugação, *-ar*, fez com que ela se tornasse o padrão de conjugação dos novos lexemas verbais, como em *deletar*, *googar*, *tuitar*. A produtividade, ainda, conduz a uma maior aceitabilidade de padrões frequentes em julgamentos de gramaticalidade e de similaridade do que dos padrões pouco produtivos, já que os indivíduos têm consciência da frequência fonotática dos padrões linguísticos (cf. VITEVITCH et al, 1997; 1999; BYBEE, 2010).

A produtividade dos padrões fonológicos, que é expressa por meio da frequência de tipo, torna-se um gatilho que desencadeia os processos analógicos, os quais são frutos de processos cognitivos de comparação por extensão. Dessa maneira, extensão analógica refere-se à extensão de um padrão fonológico frequente de um vocábulo a outro. Quanto maior a similaridade e menor a frequência de um dos itens, maior a probabilidade de que ocorra extensão analógica do vocábulo mais frequente para o menos frequente (cf. BYBEE, 2010). Tanto a frequência de tipo quanto a extensão analógica consistem em conceitos essenciais para o desenvolvimento desta pesquisa, posto que as predições teóricas e as decisões metodológicas pautam-se nos efeitos de ambas.

Os efeitos de frequência são atestados em diversos estudos de Linguística e de Psicologia (cf. VITEVITCH et al, 1997; 1999; SHADEMAN, 2006; BECKNER et al, 2009; BYBEE, 2010; CANTONI, 2013), os quais, por meio de metodologias experimentais, trazem evidências de que a frequência de uso é uma das inúmeras informações contidas nas redes, ou exemplares.

Estudos como o Vitevitch et al (1997) testaram, por meio de dois experimentos, os efeitos da probabilidade fonotática e dos padrões acentuais em pseudopalavras dissílabas no inglês. No primeiro experimento, os falantes julgavam a gramaticalidade das sequências sonoras, as quais alternavam entre

padrões de maior e menor probabilidade fonológica com diferentes padrões acentuais. Os resultados demonstraram que os falantes apresentavam intuições a respeito da probabilidade fonotática das sequências sonoras, isso porque eles julgavam os padrões fonológicos com maiores índices de frequência como melhores do que as sequências sonoras com menores índices de frequência fonotática. No segundo experimento, os autores verificaram se o tempo de processamento dos padrões fonológicos seguiria a sua probabilidade fonotática. Novamente, os falantes reagiram mais rapidamente às sequências sonoras com padrões fonológicos mais frequentes do que com os menos frequentes. Esses resultados evidenciam que a frequência fonotática dos padrões fonológicos é importante para o componente fonológico e para o acesso lexical.

Ademais, Shademan (2006) testou como a analogia lexical (ou similaridade lexical¹²) e o conhecimento gramatical (probabilidade fonotática) podem ter sua importância alterada de acordo com o desenho experimental e os estímulos empregados. No primeiro grupo experimental, a autora inquiriu, por meio de testes de gramaticalidade, a aceitabilidade de vocábulos novos, os quais obedeciam a métricas específicas que controlavam a probabilidade fonotática e a similaridade lexical. Os resultados apontaram que a probabilidade fonotática apresentava efeitos mais fortes do que a similaridade lexical, embora ambas tenham se mostrado significativas. O segundo grupo experimental, também por meio de julgamento de gramaticalidade, testava os efeitos de similaridade lexical e probabilidade fonotática por meio de estímulos com palavras novas e palavras reais. Os resultados demonstraram predominância de efeitos de similaridade lexical sobre a probabilidade fonotática. Isto é, os efeitos de similaridade lexical e de probabilidade fonotática podem ter sua significância alterada conforme os estímulos utilizados. De maneira geral, o estudo de Shademan (2006) mostra que os efeitos de similaridade lexical (por analogia) manifestam-se em condições experimentais que contemplam tanto vocábulos novos quanto vocábulos reais, sendo mais acentuados quando compreendiam ambos.

Os estudos de Vitevitch et al (1997) e de Shademan (2006) são impor-

¹² Similaridade lexical consiste em relações de semelhanças fonológicas entre vocábulos (no caso, pseudo-palavras e palavras reais).

tantes para demonstrar que a frequência dos padrões fonológicos impacta na gramática fonológica do falante de modo a predizer as suas escolhas e os seus julgamentos de gramaticalidade. Além disso, eles demonstram que os efeitos de frequência e que os processos analógicos também ocorrem em estudos que se valem de não palavras, ou pseudopalavras. Por mais que pseudopalavras não possuam frequência de uso, o seu estudo permite atestar que os falantes possuem gramática fonológica e, por isso, eles podem produzir, e produzem, vocábulos desconhecidos ou inexistentes em sua língua materna. Os esquemas abstratos permitem que os padrões fonológicos emergjam, preenchendo, através de extensões analógicas, as lacunas distribucionais.

Nos estudos linguísticos, a frequência de tipo e de ocorrência é checada em *corpora* de fala ou escrita, ou em ambos. A checagem dos índices de frequência é realizada a fim de determinar a força das conexões fonológicas e os efeitos de frequência, os quais são importantes para fazer predições a respeito do comportamento fonológico dos falantes. Este estudo, ao se valer de pseudopalavras, objetiva investigar as tendências que a gramática fonológica apresenta no que se refere aos tipos silábicos e ao sistema acentual do PB a partir da frequência de tipo. Para esta dissertação, a frequência de tipo e a extensão analógica mostram-se essenciais para as predições teóricas e as decisões metodológicas assumidas adiante.

1.5 Considerações Finais

Este capítulo apresentou o percurso teórico que a presente pesquisa realizou. Na exposição das características fonéticas do acento primário, verificou-se a duração como o principal correlato acústico para o acento em PB, o qual foi utilizado como parâmetro para a codificação dos dados desta pesquisa. A segunda parte deste capítulo dedicou-se à exposição das implicações da relação entre sílaba e acento, apresentando para tal duas propostas de sílaba: a de Selkirk (1982) e a de Stuart (1982). Em sequência, resenharam-se as principais análises dedicadas ao acento primário no PB com a finalidade de evidenciar os méritos e as lacunas encontradas em cada uma delas. Ainda, expôs-se a perspectiva de análise a partir da qual esta pesquisa partiu, a

saber, a abordagem Multirrepresentacional.

Ao considerar uma perspectiva de análise para o acento primário em PB pautada no uso real e ao ter como objetivo a eliciação de dados empíricos, apresentaram-se, na quinta parte deste capítulo, os pressupostos teóricos dos Modelos Multirrepresentacionais, ressaltando a inter-relação entre processos cognitivos e experiência de uso a fim de explicitar a concepção de língua que orientou esta dissertação.

2

Metodologia

Este capítulo apresenta as decisões metodológicas adotadas no desenho experimental desta pesquisa, a qual investigou em que medida os tipos silábicos contribuem para a determinação do acento primário do português brasileiro. Para isso, enfrentaram-se, previamente, desafios metodológicos vultosos, os quais serão discutidos em detalhes na seção 2.1. Em sequência, explicita-se, na seção 2.2, a metodologia final do experimento realizado, apresentando, para isso, na seção 2.2.1, as informações relativas ao perfil dos indivíduos que participaram do experimento, bem como, na seção 2.2.2, justificam-se as motivações que levaram à seleção dos tipos silábicos alvos desta investigação e a forma de composição dos estímulos, seguida, na seção 2.2.3, da quantidade de dados elicitados. Ainda, na seção 2.2.4, apresentam-se os procedimentos adotados durante a realização do experimento e, na seção 2.3, expõem-se o modo de codificação dos dados, as variáveis selecionadas e os modelos estatísticos utilizados na análise dos dados. Por fim, na seção 2.4, realizam-se as considerações finais deste capítulo.

2.1 Desafios Metodológicos

Esta seção apresenta e discute os desafios metodológicos que esta pesquisa enfrentou a fim de investigar o acento primário no PB. Para isso, mostra-se, na seção 2.1.1, como e em quais contextos diacríticos são utilizados no sistema ortográfico do português para codificar o acento prosódico. Em sequência, na

seção 2.1.2, apresenta-se e discute-se como o conhecimento ortográfico dos falantes, proveniente da escolarização, pode influir em tarefas linguísticas, demonstrando, com isso, as implicações e as limitações que a utilização de estímulos ortográficos traria para este estudo, visto que, como dito, o português codifica, em alguns contextos, o acento prosódico por meio do acento gráfico. Estas duas seções são essenciais para motivar a realização de um metaexperimento, o qual consiste em um experimento que verifica a própria condição experimental, exposto na seção 2.1.3. Por fim, na seção 2.1.4, explicitam-se as motivações que levaram à compilação de um *corpus* linguístico, denominado de *Corpus* ABG, com a finalidade de ser fonte de dados para a extração de frequência e dos padrões fonológicos do PB. Vale a ressalva de que todos esses passos consistem em decisões metodológicas que tiveram que ser tomadas previamente à realização do experimento deste estudo.

2.1.1 A Ortografia

O sistema de escrita da Língua Portuguesa consiste em um sistema alfabético derivado do alfabeto latino, o qual se utiliza de sinais de acentuação gráfica para “amplia[r] a capacidade de representar a tonicidade, sem recorrer à inclusão de novos grafemas que descaracterizariam o alfabeto latino e o tornariam mais complexo” (MARRA, 2012, p. 15).

A amplificação do sistema alfabético foi realizada com o acréscimo de sinais auxiliares, ou *notações léxicas*, como denominam as gramáticas normativas (cf. CAVALIERE, 2009; CUNHA & CINTRA, 2008), os quais objetivam indicar a exata pronúncia das palavras. As notações consistem em: acento agudo e circunflexo, til, trema, apóstrofo, cedilha e hífen.

Entre esses diacríticos, o acento gráfico é o único que tem como finalidade designar a sílaba mais proeminente de algumas palavras. O critério para determinar quais vocábulos devem receber a marcação gráfica está pautado na concepção linguística de acento marcado. Independentemente do mecanismo que possa reger a atribuição acentual, pode-se dizer que a literatura linguística para o acento primário no português distingue o acento prosódico dentro de

duas categorias: o acento regular e o acento irregular (cf. MARRA, 2012; NEY, 2014).

O acento "regular" compreende as oxítonas terminadas em consoante, *colher* e *civil*, e as paroxítonas terminadas em vogal, *menino* e *gato*¹. Já o acento "irregular", que recebe diacrítico, é composto pelas proparoxítonas, *abóbora* e *médico*, pelas paroxítonas com sílabas finais pesadas, *açúcar* e *útil*, e pelas oxítonas com sílabas finais leves, *café* e *cajá*. Além desses padrões acentuais, há os seguintes casos: *i* e *u* em posição final com acento paroxítono, seguidos ou não de *s*, *júri* e *bônus*; os ditongos semiabertos com acento oxítono, *herói* e *chapéu*; hiatos *i* e *u*, *saúde* e *país*; ênclise e mesóclise de pronomes átonos com formas verbais, *movê-las* e *trá-lo-ás*; etc (cf. CUNHA & CINTRA, 2008). Como pode ser observado, apenas o segundo grupo recebe a marcação gráfica indicativa de tonicidade, a qual tem como propósito alertar o falante de que a palavra não possui o acento regular, ou não marcado, da Língua Portuguesa.

Os sinais utilizados para a marcação gráfica podem ser o acento agudo ou o circunflexo. O acento agudo codifica a abertura das vogais tônicas *a*, *e* e *o*, como em *amável*, *exército* e *herói*, e o fechamento das vogais tônicas *i* e *u*, como em *horrível* e *açúcar*, enquanto o circunflexo indica o fechamento das vogais tônicas *a*, *e* e *o*, como *câmara*, *mês* e *avô* (cf. BECHARA, 2004; CUNHA & CINTRA, 2008).

A codificação gráfica foi um recurso estabelecido na Língua Portuguesa apenas no início do século XX, visto que, segundo Marra (2012), a escrita portuguesa passou por três fases até que o sistema ortográfico começasse a ser sistematizado seguindo princípios linguísticos. As fases, segundo o autor, consistem em: i. Ortográfica (XII-XVI); ii. Pseudoetimológica (XVI-XX); e iii. Contemporânea (XX-atual)². A partir deste período, diversos acordos ortográficos foram publicados, os quais buscavam sistematizar as variedades linguísticas, sobretudo entre Portugal e Brasil, em prol de uma ortografia em comum. Foram essas reformulações que resultaram em um sistema ortográfico

¹ Note que o acento "regular" espelha o acento não marcado da hipótese Métrica, de Bisol (1994), em que os padrões regulares, não marcados, não possuem diacrítico.

² O percurso histórico da acentuação gráfica bem como as Reformas Ortográficas estabelecidas desde o princípio do século XX podem ser verificados em detalhes em Marra (2012).

que assinala o acento prosódico de palavras marcadas por meio de um acento gráfico.

2.1.2 Conhecimento Ortográfico

Apesar de a literatura a respeito do processamento de leitura e de palavras não ser consensual e unânime, a visão adotada neste trabalho assume os pressupostos de Mattingly (1984; 1992), para o qual há dois processos mentais que subdividem as atividades linguísticas. O primeiro, denominado de atividades linguísticas primárias, inclui a fala e a audição, as quais são naturais a todos os seres humanos. O segundo, chamado de atividades linguísticas secundárias, requer a consciência linguística ou metalinguística que é proveniente de aspectos culturais do indivíduo, tais como a versificação e a leitura.

Segundo o autor,

todos os usuários da língua normal, como tem sido afirmado, têm acesso aos conteúdos das representações linguísticas. Isto significa que eles têm uma habilidade potencial para introspecção e para reportar detalhes significativos da representação, e para considerá-la como uma estrutura de frases, palavras e segmentos, não que eles realmente podem fazê-lo (...) Consciência linguística deve ser, em grande parte, adquirida. O principal estímulo para a consciência linguística nas culturas modernas é a alfabetização³ (MATTINGLY, 1992, p. 15).

De modo semelhante ao aprendizado da escrita, a consciência linguística aumentaria e se diversificaria a partir do contato com a ortografia. Para Mattingly, a consciência linguística do indivíduo não está no entendimento de como ocorre o processo de produção e de compreensão das sentenças, mas sim no conhecimento da estrutura gramatical delas. "O estado de se ter acesso ao

³ Trecho Original: "All normal language users, it has been claimed, have access to the contents of linguistic representations. This means that they have a potential ability to introspect and report on significant details of the representation, and to regard it as a structure of phrases, words, and segments, not that they can actually do so (...) Linguistic awareness must in large part be acquired. The principal stimulus for linguistic awareness in modern cultures is literacy".

conhecimento gramatical é o que eu quis dizer por consciência linguística"⁴ (MATTINGLY, 1984, p. 18). Para o autor, essa consciência caracteriza-se por ser um processo eminentemente natural, uma vez que é um estado mental que se assemelha ao dos aprendizes de línguas.

No que tange ao sistema acentual, o nível de escolaridade e a qualidade no processo de alfabetização dos indivíduos são os fatores que parecem motivar a consciência linguística. De modo geral, sujeitos com Ensino Superior tendem a ter maiores percepções a respeito da marcação do acento prosódico por meio do acento gráfico. Em compensação, os falantes que possuem apenas o Ensino Médio, ou mesmo aqueles que não possuem um grande domínio das estruturas gramaticais, têm menos consciência da função do acento gráfico, bem como da sílaba mais proeminente nos vocábulos.

Inúmeros exemplos podem ser ilustrativos de como os falantes nem sempre têm consciência da relação entre acento e ortografia. O caso mais saliente corresponde às palavras que sofrem processos de flutuação acentual (cf. BENEVIDES, 2013; 2015), ou *silabada* (cf. LIMA, 2002; BECHARA, 2005; CUNHA & CINTRA, 2008). Essas palavras são, frequentemente, produzidas com o acento oscilante em detrimento do acento regular, como em *ruim/ruim*, *tríplex/tríplex* e *ureter/ureter*. Além delas, é possível encontrar casos em que, apesar de a escrita indicar e determinar um padrão acentual, o uso cristalizou outro padrão, sem que haja qualquer indicativo de consciência dessa modificação, como em *Janaina*, [ʒa'na.j.nɐ] > [ʒa.na'i.nɐ], e *Ramila*, [xa'mi.lɐ] > ['xa.mi.lɐ].

Os nomes próprios são ótimos exemplos para mostrar que o arraigamento do acento prosódico faz com que, muitas vezes, os falantes não percebam ou ignorem a marcação ortográfica. A grafia dos nomes de pessoas e de estabelecimentos demonstra como a "opcionalidade" do diacrítico em nada altera a sua produção, como *Luís/Luis*, *Antônio/Antonio*, *Lívia/Livia*, *Essencial/Essencia*⁵.

⁴ Trecho Original: "The state of having access to one's grammatical knowledge is what I meant by linguistic awareness".

⁵ O estabelecimento ao atribuir esta grafia, provavelmente, teve a intenção de aproximar a grafia à pronúncia do vocábulo em inglês, contudo, tanto os clientes quanto os funcionários produzem como se não houvesse a marcação gráfica.

Apesar da aparente inconsciência ortográfica, a literatura apresenta inúmeros estudos que evidenciam o uso de informações ortográficas nas tarefas linguísticas. Seidenberg & Tanenhaus (1979) investigaram, por meio de tarefas de *priming*⁶, se mapeamentos puramente acústicos poderiam ser visualizados em tarefas de monitoramento de rimas com estímulos visuais e auditivos. Nesse estudo, os indivíduos deveriam determinar semelhanças ou dessemelhanças entre os pares de palavra alvo-pista, como em *pie-tie* e *rye-tie*, respectivamente. Os resultados demonstraram que os falantes fazem uso de informações ortográficas e acústicas na tentativa de mapear as palavras-alvo em ambas as direções. Os autores ainda afirmam que, em sujeitos alfabetizados, as informações ortográficas tornam-se disponíveis automaticamente, bem como as fonológicas.

Além desses autores, o estudo de Castro & Morais (não publicado apud MORAIS et al, 1998), realizado com falantes alfabetizados e analfabetos do português de Portugal, por meio de dois estímulos auditivos, investigou o processo de fusão fonológica em palavras com consistência ortográfica, como em *cara-lara*, tendo como alvo a palavra *clara*, e em palavras com inconsistência ortográfica, como em *liz-fiz*, tendo como alvo *feliz* (no português europeu, a vogal é /ə/ e a pronúncia é /fliʃ/⁷). Os resultados mostraram que tanto os falantes alfabetizados como os analfabetos realizavam fusões fonológicas em palavras com consistência ortográfica, entretanto, somente os analfabetos produziam fusões fonológicas em palavras com inconsistência ortográfica. Tais dados sugerem que o conhecimento ortográfico dos sujeitos alfabetizados inibe a realização das fusões fonéticas.

Por fim, o único estudo encontrado que investiga a relação entre acento lexical e pistas ortográficas é o de Kelly, Morris & Verrekia (1998) na língua inglesa. Diferentemente do português, o inglês não apresenta um acento

⁶ Segundo França et al. (2008, p. 35), "o paradigma de *priming* no acesso lexical é um teste para desvendar aspectos da arquitetura do léxico mental, por exemplo, os critérios de agrupamento. Agrupam-se palavras inteiras ou fatias morfológicas? Por semelhança fonológica ou semântica? Para isso, os experimentadores constroem estímulos de muitos pares de palavras. Na metade dos casos, a primeira palavra do par, tecnicamente chamada de *prime*, terá algum relacionamento com a segunda palavra do par, tecnicamente chamada de alvo; por exemplo, no par *caro-carinho* há uma relação de semelhança fonológica entre *prime* e alvo".

⁷ Ressalta-se que, embora a transcrição corrente da vogal alta anterior centralizada seja [i] para o português europeu, a vogal é transcrita como [ɨ] pelos autores.

gráfico que codifica a tonicidade das palavras, em contrapartida, segundo os autores, "o inglês contém muitas outras pistas ortográficas de acento cuja preditividade pode ser verificada por meio da análise do léxico do inglês" (p. 822)⁸.

O estudo em questão, por meio de tarefas de nomeação e de decisão lexical com palavras dissílabas, verificou que pistas segmentais são utilizadas para codificar o acento na segunda sílaba. Isso porque mais letras são usadas para marcar as informações fonêmicas do acento iâmbico, ou marcado, como em *discuss*, *roulette* e *dinette*, enquanto o acento trocaico, ou não marcado, não requer essas informações, como em *discus*, *comet* e *sonnet*.

Os autores concluíram que "embora poucos falantes alfabetizados tenham consciência dessas relações entre acento e ortografia, pesquisas têm demonstrado que eles têm conhecimento implícito delas" (KELLY, MORRIS & VERREKIA, 1998, p. 823)⁹. Sendo assim, "falantes alfabetizados do inglês não somente usam a ortografia para determinar as representações fonêmicas de palavras escritas, mas também para determinar seus padrões acentuais" (KELLY, MORRIS & VERREKIA, 1998, p. 825)¹⁰.

Diante do exposto, constatou-se que o conhecimento de regras ortográficas tende a ser ativado automaticamente em falantes alfabetizados e que esse conhecimento parece ser inerente e implicitamente utilizado em tarefas de nomeação mesmo em sistemas linguísticos em que o acento gráfico não é empregado para codificar o acento prosódico (cf. KELLY, MORRIS & VERREKIA, 1998; SEIDENBERG & TANENHAUS, 1979). Tais fatos conduzem à pressuposição de que, por o sistema ortográfico do português conter diacríticos que codificam e indicam o acento prosódico, os falantes tenderiam a ser mais conscientes das regras ortográficas do que os falantes do inglês.

As considerações apresentadas acima resultam na problemática de encontrar um paradigma experimental que permita diagnosticar os processos

⁸ Trecho Original: "English contains many other orthographic cues to stress whose predictiveness can be verified through analyses of the English lexicon".

⁹ Trecho Original: "although few literate speakers are consciously aware of these relations between stress and spelling, research has shown that they have implicit knowledge of them".

¹⁰ Trecho Original: "literate English speakers not only use orthography to determine the phonemic representations of written words but also to determine their stress patterns".

fonológicos empregados por falantes nativos do português brasileiro na atribuição acentual. Isso porque desenhos experimentais pautados em estímulos ortográficos, possivelmente, direcionam a produção dos vocábulos às regras de acentuação gráfica. Ressalta-se que tanto a presença quanto a ausência da marca gráfica indicam aos indivíduos a sílaba mais proeminente do vocábulo, como em *sábia*, *sabia* e *sabiá*. O emprego do diacrítico, em *sábia* e *sabiá*, mostra que os vocábulos devem ser produzidos com tonicidade proparoxítona e oxítona, respectivamente; em contrapartida, a sua ausência, em *sabia*, revela que a tonicidade desejada consiste na paroxítona.

Apesar dos indícios de consciência ortográfica, rejeitar, *a priori*, o uso de estímulos ortográficos sem realizar qualquer teste a fim de verificar possíveis enviesamentos experimentais consistiria numa decisão precipitada, posto que a sua utilização traria ganhos substanciais em termos de tempo de confecção do experimento e de sua aplicação. Por essa razão, decidiu-se realizar um estudo-piloto em que o modo de apresentação dos estímulos fosse o objeto de investigação, o qual passa a ser denominado de metaexperimento, a ser descrito a seguir.

2.1.3 Metaexperimento

A seleção das formas de exibição dos estímulos aos participantes pautou-se nos estudos da Fonologia de Laboratório (cf. COHN, FOUGERON & HUFFMAN, 2012). As condições experimentais consistiram em áudio, imagem e ortografia. A metodologia adotada para cada condição bem como uma síntese dos resultados gerais serão apresentadas adiante.

Condição Áudio

Na condição áudio, os estímulos foram apresentados por meio de áudios, os quais consistiam em produções isoladas das sílabas de cada palavra ou pseudopalavra, apresentando características de sílaba tônica, segundo o parâmetro de duração (cf. ESCUDERO et al, 2009). Cada sílaba foi inserida numa rotina de três sílabas sequenciais, com intervalo de 0,5s entre cada sílaba e 2s entre as palavras. Após a apresentação do áudio das três sílabas, o informante

era instruído a produzir a palavra formada mantendo a mesma ordem de apresentação das sílabas. Quando o indivíduo produzia a palavra de modo silabificado, a pesquisadora solicitava sua repetição.

Condição Imagem

Segundo Schiller (2012),

(...) uma das tarefas mais desafiadora dos pesquisadores de produção de linguagem é desenhar paradigmas experimentais para elicitare palavras-alvo de uma forma controlada, sem apresentar o próprio estímulo verbal a ser elicitado aos participantes - caso contrário, a compreensão da linguagem está envolvida também. Muitas vezes, isso é feito pelos participantes por meio da nomeação de imagens. As imagens são estímulos não linguísticos, as quais podem ser facilmente reconhecidas pelo sistema visual-perceptual e, posteriormente, nomeadas. Supõe-se que, de uma maneira geral, após o reconhecimento visual de um estímulo imagético, este estímulo conduz à conceituação do significado da imagem, que depois pode ser verbalmente codificado por meio da ativação e recuperação do seu item lexical correspondente no léxico mental do participante (SCHILLER, 2012, p. 565-66)¹¹.

Tendo em vista esta afirmativa de Schiller (2012) e os estudos de Veloso (2003), decidiu-se empregar imagens como forma de elicitare as pseudopalavras. Entretanto, como elas não dispõem de imagens que possam representá-las, já que são palavras inventadas, resolveu-se, em analogia às brincadeiras de adivinhação realizadas com figuras, utilizar imagens prototípicas como forma de motivar a composição das palavras deste experimento.

¹¹ Trecho Original: "(...) one of the most challenging tasks of language production researchers is to design experimental paradigms to elicit target words in a controlled way without presenting the to-be-elicited verbal stimulus itself to participants – otherwise language comprehension is involved as well. Often, this is done by having participants name pictures. Pictures are non-linguistic stimuli, which can easily be recognized by the visual-perceptual system and subsequently named. It is generally assumed that after the visual recognition of a pictorial stimulus, this stimulus leads to the conceptualization of the meaning of the picture, which can afterwards be verbally encoded by activating and retrieving its corresponding lexical item from the participant's mental lexicon".

A tarefa consistia na extração da primeira sílaba da palavra que nomeava a imagem exibida, ou seja, considerando as imagens da Figura 2.1, **CI** de *cigarro*, **TA** de *tatuagem*, **BA** de *bateria*. Após as figuras, um *slide* com a pergunta “Qual palavra foi formada?” era apresentado. Nesse momento, o participante deveria produzir a palavra formada, no caso, [si.ta.ba].

Figura 2.1 Exemplo de imagens prototípicas empregadas para incitar a pseudopalavra [si.ta.ba].

Para assegurar que os informantes atribuiriam o mesmo nome às imagens utilizadas, realizou-se uma fase de reconhecimento de imagens. Nela, todas as 57 imagens foram exibidas sequencialmente com seu respectivo nome, apresentado de forma escrita. Segundo Schiller (2012), essa estratégia tem a finalidade de potencializar o reconhecimento preciso das imagens. Após a fase de reconhecimento de imagens, as instruções para a execução do teste eram fornecidas aos informantes e, para que eles compreendessem a tarefa, uma fase de treino foi realizada.

Condição Ortografia

Os estímulos foram exibidos por meio de frases, as quais, apesar de não terem a estrutura fixa de uma frase veículo, possuíam características semelhantes: [Sujeito + Verbo + Objeto Direto + Objeto Indireto]. A pseudopalavra foi inserida em posição de objeto direto a fim de evitar que os falantes atribuísem interpretações verbais, posto que, como as teorias fonológicas predizem (cf. LEE, 1995; MATEUS & D'ANDRADE, 2002), verbos e não

verbos comportam-se de modo distinto.

Também por questões metodológicas, determinou-se que nenhum estímulo seria apresentado com marcação de acento gráfico, mesmo que a Norma Padrão da Língua Portuguesa assim exigisse. Isso porque foi tomado como premissa que os falantes alfabetizados poderiam ter maior consciência das regras ortográficas de acentuação quando o acento gráfico estivesse presente. A decisão pautou-se nos estudos de Benevides (2013), a qual apresentou pseudopalavras repetidas três vezes ao longo de um texto, sendo que a pseudopalavra ortográfica era exibida de três modos distintos, isto é, com marcação gráfica para atribuir acento oxítono, paroxítono e proparoxítono.

No estudo, verificou-se que os falantes com nível de escolaridade correspondente ao Ensino Médio tendiam a produzir as pseudopalavras com a mesma tonicidade nas três ocorrências, independentemente da marcação gráfica da pseudopalavra ortográfica. A tonicidade atribuída na primeira produção tendia a ser repetida em todas as demais e, ainda, não necessariamente a tonicidade empregada na primeira produção correspondia à tonicidade que as regras ortográficas determinavam. Em contrapartida, os falantes com Ensino Superior pareciam ser mais conscientes da distinção, aplicando, sobretudo os informantes com faixas etárias mais altas, a tonicidade determinada pela marcação gráfica.

Em decorrência disso, todos os vocábulos, reais ou não reais, foram apresentados sem qualquer marcação gráfica, como em “*Marcia comprou o movel da cozinha*”. Para isso, realizou-se uma fase de treino em que palavras reais e de alta frequência de ocorrência (cf. ASPA, 2014), as quais deveriam ter acento gráfico segundo as Normas Ortográficas da Língua Portuguesa, foram exibidas sem acento gráfico. Esperava-se, assim, que o falante se habituassem a ler naturalmente as palavras reais e aplicasse semelhante comportamento nas pseudopalavras.

Participantes

O teste-piloto foi realizado com 9 falantes nativos do português brasileiro, 4 homens e 5 mulheres, com idade entre 20 e 50 anos, 6 falantes com Ensino

Superior e 3 com Ensino Médio - um em cada condição experimental. Todos possuem audição e visão normais e foram participantes voluntários. Cada condição experimental contou com 3 informantes.

Estímulos

Os estímulos compreenderam 5 estruturas de palavras¹² com 4 pseudopalavras cada, totalizando 20 pseudopalavras. Somaram-se a eles, 20 palavras reais com estruturas silábicas semelhantes às pseudopalavras, as quais tinham a função de ser palavras distratoras. Todos os estímulos, alvo e distrator, foram aleatorizados por meio da função *aleatório entre* do Microsoft Excel, com a mesma sequência para todos os sujeitos. A fim de evitar habituação, os estímulos foram repetidos 3 vezes, totalizando 120 produções por indivíduo, 360 por condição experimental e 1.080 em conjunto.

Resultados

Os resultados gerais de atribuição acentual no metaexperimento correspondem a 7 (1,3%) ocorrências de proparoxítonas, 275 (52%) de paroxítonas e 247 (46,7%) de oxítonas, com $\chi^2 = 246.14$ (2), $p < 0.0001$, os quais podem ser visualizados por *condição experimental*¹³ na Tabela 2.1¹⁴.

Tabela 2.1 Atribuição de *tonicidade* por *condição experimental* no metaexperimento¹⁵.

Condição Experimental	Proparoxítona	Paroxítona	Oxítona
Áudio	0	75	95
Imagem	7	104	69
Ortografia	0	96	83
Total	7 (1,3%)	275 (52%)	247 (46,7%)

¹² As estruturas de palavras foram CV-CV-CV, CV-CV-CVC, CV-CVC-CVC, CV-CVC-CV e CVC-CV-CV.

¹³ Por fins de clareza, as variáveis serão apresentadas em *itálico*.

¹⁴ Veja os resultados completos por *pseudopalavra* e *condição experimental* no anexo A.

¹⁵ A quantidade de estímulos por condição experimental diverge em decorrência de dados que foram descartados por silabificação, inserção ou apagamento de segmentos.

As análises estatísticas¹⁶ realizadas por meio de modelos de regressão logística binomial constataram a significância da variável *condição experimental*, com as condições imagem e ortografia favorecendo a tonicidade paroxítona e a condição áudio, a oxítona¹⁷.

O favorecimento do acento oxítono na condição áudio é evidenciada já nos dados percentuais com 55,9% de atribuições oxítonas e 44,1% de paroxítonas (veja o anexo A). A diferença entre as tonicidades, além de contrariar as previsões da literatura a respeito da predominância do acento paroxítono sobre os demais padrões, não se mostra significativa, já que obteve $\chi^2 = 2.3529$ (1) e $p = 0.125$. A ausência de significância na atribuição acentual faz com que a hipótese nula (H^0) não possa ser rejeitada, isto é, não foi possível demonstrar diferença entre produções oxítonas e paroxítonas nesta condição experimental.

Os resultados inconclusivos obtidos na condição áudio podem ser devido à emergência do acento oxítono em todas as estruturas de palavras, mesmo naquelas em que as hipóteses prediriam predominância do acento paroxítono. A constância em produções acentuais oxítonas até mesmo em vocábulos como *ducirca*, *jofelta*, *zubesma*, *citaba* e *lepaque* traz indícios de enviesamento experimental, já que nas demais condições experimentais tal fato não ocorreu. Além disso, a alta taxa de dados descartados, em consequência de produções silabificadas¹⁸, 5% dos dados, complementa as razões pelas quais a condição áudio não foi considerada apropriada para a presente investigação.

Numa análise comparativa entre as condições restantes, observou-se, na condição imagem, 83 produções oxítonas e 96 paroxítonas, com $\chi^2 = 7.0809$ (1) e $p = 0.007791$. Já na condição ortografia, 69 oxítonas e 104 paroxítonas, com $\chi^2 = 0.94413$ (1) e $p = 0.3312$, o que não permitiu a rejeição da hipótese nula.

Ademais, a observação das produções acentuais por *estrutura de palavra*

¹⁶ Ressalta-se que as análises apresentadas adiante incluem somente as produções paroxítonas e oxítonas, visto que a inserção das proparoxítonas, em decorrência do seu baixo índice de produção, comprometeria a significância dos resultados.

¹⁷ O valor de *intercept*, respectivamente, corresponde a 1.8366, 0.9111 e -0.2337, sendo que só os dois primeiros são significativos.

¹⁸ Apesar da pesquisadora ter solicitado a repetição dos estímulos, os falantes continuaram a realizar produções silabificadas.

e *pseudopalavra* indica que o conhecimento de regras ortográficas pode ter direcionado a atribuição acentual em alguns vocábulos da condição ortografia. Essa afirmativa pauta-se nas produções destoantes - em relação à condição imagem - nas estruturas de palavra CV-CV-CV e CVC-CV-CV (veja o anexo A). Na primeira, das 4 pseudopalavras, apenas uma, a saber, *casemi*, terminava com *-i*, a qual, segundo as regras ortográficas da Língua Portuguesa (cf. BECHARA, 2004; CUNHA & CINTRA, 2008), na ausência de acento gráfico, a sílaba final ([mi]) deve ser tônica, sendo esse o único vocábulo a receber, categoricamente, acento oxítono. Na segunda, das 4 pseudopalavras, 3 terminavam em *-i* ou *-u*, as quais, novamente, foram as únicas a receber acento oxítono. Os resultados indicam, portanto, que a condição ortografia está sujeita a interferências provenientes do conhecimento ortográfico dos falantes, como apontaram os estudos de Seidenberg & Tanenhaus (1979), Castro & Moraes (não publicado apud MORAIS et al, 1998) e Kelly, Morris & Verrekia (1998).

Observou-se, portanto, que tanto a condição áudio como a condição ortografia não se mostram adequadas para a realização deste estudo, na medida em que podem sofrer interferências da condição experimental, ou mesmo de conhecimentos ortográficos. Deve-se às constatações apresentadas acima que esta pesquisa assume a condição imagem como a mais adequada para eliciar o acento primário em pseudopalavras no PB.

2.1.4 *Corpus* ABG

O *Corpus* ABG consiste em um *corpus* linguístico compilado com o objetivo de ser fonte de dados deste estudo¹⁹. Diante da diversidade de *corpora* existentes para o português, a decisão de compilar um *corpus* para ser base de dados desta pesquisa deveu-se à ausência de um *corpus* que fornecesse os dados necessários para a computação da frequência por tipos silábicos e que ainda permitisse acessar aos dados para a realização das mais variadas buscas, como a de estrutura da palavra e de sílaba, de categoria gramatical.

¹⁹ O *Corpus* ABG, bem como as ferramentas computacionais a serem descritas, foi realizado em parceria com Bruno Ferrari Guide, sob orientação do Prof. Dr. Marcelo Barra Ferreira.

A compilação foi realizada a partir dos pressupostos da Linguística de *Corpus*, a qual se caracteriza por ser uma metodologia que determina o modo de aplicação de um conjunto de pressupostos na formação de *corpora* linguísticos. Seu instrumental tem bases empíricas e probabilísticas, já que os seus dados provêm de usos concretos da língua, compreendendo a linguagem como um sistema probabilístico (cf. SARDINHA, 2000a). Assumiu-se, como *corpus*, a seguinte definição de Sanchez (1995, p. 8-9 apud SARDINHA, 2000a, p. 338):

Um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise.

O excerto acima apresenta os conceitos primordiais que subjazem à Linguística de *Corpus*, os quais consistem em: autenticidade, especificidade, adequação, extensão e representatividade, conceitos esses que foram seguidos no processo de compilação do *Corpus* ABG.

O *corpus* é composto de textos escritos e orais produzidos por falantes nativos do português brasileiro, os quais compreendem, no *corpus* de fala, transcrições ortográficas de entrevistas Sociolinguísticas do Projeto SP2010 (MENDES, 2010), C-ORAL-BRASIL (RASO & MELLO, 2012), Projeto SP2010 (MENDES & OUSHIRO, 2012) e Iboruna (GONÇALVES, 2014), e, no *corpus* escrito, artigos científicos, textos jornalísticos e postagem de blogs.

A seleção desses textos contabilizou 3.616.625 ocorrências de palavras e 92.602 tipos de palavras, sendo que 1.938.805 ocorrências são provenientes dos *corpora* de fala e 1.676.820, dos *corpora* escritos (veja a Tabela 2.2). A quantidade de palavras o caracteriza como um *corpus* médio-grande, segundo Sardinha (2000b, p. 346)²⁰.

²⁰ Os dados do *corpus* serão disponibilizados em trabalhos futuros.

Tabela 2.2 Frequência de ocorrência e de tipo por *corpus*.

<i>Corpus</i>	Frequência de Ocorrência	Frequência de Tipo
<i>Corpus</i> de Fala	1.938.805	36.484
<i>Corpus</i> Escrito	1.676.820	77.669
<i>Corpus</i> Geral	3.615.625	92.602

Tendo em vista que o *corpus* foi compilado com o objetivo de extrair dados fonológicos a respeito do acento primário no português brasileiro, a necessidade de buscar padrões fez com que algumas ferramentas fossem desenvolvidas e utilizadas por meios computacionais. Para tanto, utilizou-se um silabificador fonológico²¹, o qual tem como *input* a palavra ortográfica e como *output* a palavra silabificada. Em sequência, desenvolveu-se, especificamente para o *Corpus* ABG, um transcritor fonológico, o qual tem como *input* a palavra silabificada e como *output* a palavra silabificada e transcrita. Para que a marcação acentual pudesse ser realizada, utilizou-se o categorizador morfológico/lematizador *TreeTagger* (SCHMID, 2015), o qual consiste numa ferramenta de anotação de texto que fornece a palavra, a categoria morfológica e o lema. Ao ter a informação da categoria gramatical da palavra, desenvolveu-se um acentuador, o qual, a partir de informações ortográficas e da categoria da palavra, acentuava o vocábulo seguindo a proposta de peso silábico de Bisol (1994)²². Por fim, desenvolveu-se um codificador de estruturas fonológicas, o qual codificava as vogais, consoantes e glides em, respectivamente, V, C e G. Com todas essas ferramentas e informações, foi possível tecer diferentes buscas e quantificações²³.

Como o objeto teórico desta pesquisa é o acento primário no português brasileiro, considerando para isso os padrões que emergem na língua, o

²¹ Desenvolvido pelo Prof. Dr. Marcelo Barra Ferreira e, gentilmente, fornecido para a realização deste trabalho.

²² Ao se valer das informações ortográficas, o programa assinalava a sílaba tônica de todas as palavras que possuíam acento gráfico e, posteriormente, seguia a proposta de peso silábico de Bisol (1994) para marcar a tonicidade das palavras que não tinham sinal gráfico. As regras propostas pela hipótese Métrica foram utilizadas, já que, como ressaltado na seção 2.1.1 deste capítulo, elas codificam apenas o acento irregular, marcado, por meio de diacríticos.

²³ Todos os *scripts* desenvolvidos para essa tarefa estão disponibilizados publicamente no seguinte endereço: <https://github.com/SauronGuide/corpusABG>.

primeiro dado a ser observado consiste na distribuição da tonicidade por categoria acentual, ou seja, monossílabo, oxítono, paroxítono, proparoxítono e acento na sílaba pré-antepenúltima (Tabela 2.3). Embora não seja usual, o acento pré-antepenúltimo refere-se a casos em que a epêntese da vogal [i] cria uma nova sílaba no vocábulo, fato que desloca o acento para a quarta sílaba a partir da margem direita da palavra, como em *té[kɪ]nico*, *cá[pɪ]sula*.

Tabela 2.3 Quantidade de tipos de palavras por categoria acentual.

Categoria Acentual	Quantidade de Tipos
Monossílabo	1.525
Oxítono	24.714
Paroxítono	65.511
Proparoxítono	3.797
Quarta Sílaba	40

A distribuição acentual do *Corpus* ABG assemelha-se em partes à proporção encontrada por Viaro & Guimarães-Filho (2007), veja a Tabela 2.4. A principal diferença diagnosticada entre os *corpora* consiste nos dados referentes às proparoxítonas, contudo, ela, provavelmente, é motivada pela quantidade de dados analisada e pelo tipo de *corpus*, visto que o de Viaro & Guimarães-Filho (2007) é composto de 150.875 palavras provenientes de textos escritos, os quais pertencem ao Dicionário Houaiss da Língua Portuguesa - que só contabiliza tipos e não ocorrências -, enquanto, o *Corpus* ABG tem mais de 90.000 tipos de palavras advindos de textos escritos e falados.

Tabela 2.4 Distribuição percentual do acento por categoria em Viaro & Guimarães-Filho (2007, p. 31) e no *Corpus* ABG.

	Viaro & Guimarães-Filho (2007)	<i>Corpus</i> ABG
Monossílabo	0.3% a 0.5%	1.6%
Oxítono	25%	25.9%
Paroxítono	62%	68.5%
Proparoxítono	12%	4%
Pré-Antepenúltimo	0.4%	0.04%

A determinação de uma abordagem baseada em *corpus* parte também do

pressuposto de que uma grande quantidade de dados provenientes de usos concretos da língua consiste numa base de dados palpável para a extração de (não) regularidades a respeito da língua e do fenômeno em estudo.

Quando se diz que a variação não é aleatória, na verdade, está se afirmando que a linguagem é padronizada ('patterned'). A padronização se evidencia pela recorrência, isto é, uma colocação, coligação ou estrutura, que se repete significativamente, mostra sinais de ser na verdade um padrão lexical ou léxico-gramatical. A linguagem forma padrões que apresentam regularidade (se mostram estáveis em momentos distintos, isto é, tem frequência comparável em *corpora* distintos) e variação sistemática (correlacionam-se com variedades textuais, genéricas, dialetais, etc) (SARDINHA, 2000a, p. 352).

Os padrões fonológicos encontrados no *Corpus* ABG, por exemplo, permitem com que as predições das hipóteses acentuais, descritas no capítulo anterior, sejam checadas a partir de um grande volume de dados. A hipótese Métrica, ao predizer que o acento em português é sensível ao peso silábico, do contrário, incide na penúltima sílaba, consegue explicar somente 44.390 (89,4%) dos vocábulos, os demais 5.255 (10,6%) têm que ser justificados por outros instrumentos, como a consoante final abstrata e a extrametricidade.

Para que o contraste entre a hipótese Métrica e os dados do *Corpus* ABG fique mais evidente, apresenta-se, na Tabela 2.5, a quantificação de acordo com a explicativa empregada por Bisol (1994). Verifique, em azul, os 7.274 (14,7%) vocábulos explicados pela regra de peso silábico; em verde, os 37.116 (74,7%) vocábulos justificados pelo padrão acentual; em vermelho, os 4.124 (8,3%) vocábulos em que a extrametricidade tem que atuar; e, por fim, em laranja, os 1.131 (2,3%) vocábulos em que a consoante final abstrata tem que existir. Note que a quantidade expressiva de dados incluída nos mecanismos de extrametricidade e de consoante final abstrata para que haja atribuição acentual evidencia as limitações da proposta.

Com relação à hipótese Morfológica, para a qual o acento incide na última vogal do radical, isto é, acento paroxítono em palavras que terminam com

Tabela 2.5 Tonicidade por estrutura silábica final no *Corpus* ABG.

	Proparoxítonas	Paroxítonas	Oxítonas	Total
-VC	89	751	2.447	3.287
-G(S)	4	834	4.827	5.665
-S	1.001	9.651	589	11.241
-V	1.445	27.465	542	29.452
Total	2.539	38.701	8.405	49.645

vogal temática e oxítono nos demais casos, constatou-se que, pelo menos, 4.124 (8,2%) vocábulos (proparoxítonos + paroxítonos terminados em consoante) seriam abarcados na regra de palavras marcadas. Esse número só não é maior porque, em decorrência da ausência de marcação morfológica apropriada, os dados correspondentes às paroxítonas com vogal atemática não puderam ser contabilizados. Percebe-se, a partir dessas quantificações, que a hipótese Morfológica, de maneira análoga à hipótese Métrica, atribui uma grande quantidade de dados à regra de marcação, dados que se somam às limitações observadas na proposta.

As predições da hipótese Lexical e da hipótese Multirrepresentacional não puderam ser quantificadas, posto que elas não postulam algoritmos de atribuição acentual. Pretende-se, assim, utilizar o *Corpus* ABG como fonte de frequência e inferências realizadas para os padrões fonológicos encontrados no léxico do PB para esta dissertação.

2.2 Metodologia

A partir desta seção, apresentam-se as decisões metodológicas que guiaram o desenvolvimento do experimento final desta dissertação.

2.2.1 Participantes

O experimento foi realizado com 36 falantes nativos do português brasileiro, 14 homens e 22 mulheres, de 18 a 61 anos, residentes na Região Metropolitana de São Paulo, nascidos na cidade e no interior de São Paulo, Recife e Ceará.

Todos possuíam Ensino Superior completo ou cursando, 21 de áreas diversas (Psicologia, Engenharia, Geografia) e 15 do 1º ano do curso de Letras da Universidade de São Paulo. Nenhum informante relatou qualquer problema de audição ou visão e todos foram participantes voluntários.

2.2.2 Estímulos

Os estímulos, ou pseudopalavras, foram construídos com três sílabas de extensão para que todos os padrões acentuais (proparoxítono, paroxítono e oxítono) pudessem ser produzidos. A estrutura da pseudopalavra adotada consistiu em CV-CV-CVC²⁴, a qual inclui pseudopalavras como [lo.pe.niw] e [ka.pi.gor]²⁵. Esta estrutura foi escolhida a fim de que as predições de sensibilidade ao peso silábico, proposta pela hipótese Métrica (cf. BISOL, 1994), pudessem ser verificadas e, principalmente, para que os tipos silábicos fossem analisados a partir dos pressupostos dos Modelos Multirrepresentacionais. Vale ressaltar que apenas uma estrutura foi selecionada, visto que a utilização de mais estruturas aumentaria consideravelmente a quantidade de dados a serem elicitados e analisados.

De maneira semelhante, definiram-se dois grupos de análise para compor os tipos silábicos objetos deste estudo, os quais consistem em, no grupo 1, [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], e, no grupo 2, [-or] e [-ar]. Eles foram escolhidos dentre os tipos silábicos possíveis na língua por permitirem predições de efeitos de frequência robustas, entretanto, foram, de certo modo, escolhidos de maneira aleatória, já que outros tipos silábicos também poderiam ser testados. Restringiram-se, novamente, a dois grupos em decorrência da quantidade de dados.

Após a determinação dos tipos silábicos, estes foram combinados com outras sequências sonoras permitidas pela fonotática da língua para compor a estrutura CV-CV-CVC. Ao término da composição, atentou-se para que as pseudopalavras formadas não constituíssem vizinhança fonológica com uma palavra real, por meio da permuta de apenas um fone, a fim de evitar

²⁴ Neutralizou-se aqui a distinção entre CVC e CVV. Para mais explicações, veja a seção 1.2, do Capítulo 1.

²⁵ Veja todas as pseudopalavras nas Tabelas 2.7 e 2.9.

associações analógicas diretas com outros vocábulos. A associação direta entre uma palavra real e uma pseudopalavra poderia resultar na extensão do padrão acentual de um vocábulo a outro, ao contrário do que pretende investigar este estudo - a atuação da frequência dos padrões abstratos. Além disso, cuidou-se para que os bigramas silábicos formados (i.e., [bu.da], [go.le]) fossem tipos silábicos reais da língua, na tentativa de maximizar o reconhecimento das pseudopalavras como pertencentes à Língua Portuguesa.

A criação das pseudopalavras a partir dos tipos silábicos finais teve como objetivo verificar em que medida a frequência dos tipos silábicos pode atuar na determinação da localização acentual em português. Para isso, o primeiro grupo de análise, composto por [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], inquiriu se a frequência dos tipos silábicos finais pode prever a localização acentual. O segundo grupo de análise, composto por [-or] e [-ar], investigou se a frequência dos tipos silábicos interage com efeitos de similaridade lexical. As motivações e as hipóteses que levaram à formação dos grupos de análise 1 e 2 serão apresentadas a seguir.

Grupo de Análise 1: [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw]

A escolha dos tipos silábicos [tʃiw] e [niw], que compõem o grupo de análise 1, baseou-se nos resultados obtidos no metaexperimento e no pré-teste²⁶. Neles, verificou-se que, curiosamente, a sequência fonológica [iw], em final de palavra, recebia tanto acento paroxítono quanto acento oxítono quando a consoante precedente era [t], todavia, quando a consoante que a precedia era [n], apenas o acento oxítono era produzido, como em [ver'sa.tʃiw], [su'tʃiw] e [kẽ'niw].

Diante da diversidade de comportamento, investigou-se, por meio de dados quantitativos do *Corpus* ABG, se haveria motivações provenientes da frequência de tipo nas sequências sonoras [niw] e [tʃiw]. Como demonstra a Tabela 2.6, [niw] apresenta somente tipos oxítonos, enquanto [tʃiw], paroxítono e oxítono.

Vale ressaltar que os dados da Tabela 2.6 não fazem distinção entre /Ciw/ e /Cil/. A posição adotada assume as afirmativas de Oliveira & Cristófar-

²⁶ O pré-teste objetivou avaliar o método experimental e levantar hipóteses a respeito de uma abordagem para o acento no PB a partir de tipos silábicos. Para mais detalhes, verifique o anexo C.

Tabela 2.6 Frequência de tipo das sequências [niw] e [tʃiw] por tonicidade no *Corpus* ABG.

Sequência Sonora	Paroxítona	Oxítona
[niw]	0	19
[tʃiw]	14	46

Silva (2002), para os quais, numa perspectiva sincrônica, há indícios de que os falantes apresentam a forma [w] nas representações tanto para vocábulos como *sal* quanto *degrau*. Os autores partem das formas de plural, como *sais* e *varais*, a fim de questionar os argumentos empregados a favor da representação [l] em palavras como *sal*, já que elas teriam, nas formas derivadas, como *saleiro*, a manifestação da consoante lateral. Eles demonstram que se trata de uma justificativa *ad hoc*, visto que nas formas de plural a consoante lateral não é preservada, como em *sal-sais*. Além disso, inúmeros casos de variação já neutralizam a distinção entre as formas plurais em *u* e *l*, como em *degrau-degraus-degrais*, *sal-sais-saus*, *varal-varais-varaus*. Ainda, em estudos preliminares, Oliveira & Cristófar-Silva (2002) constataram produções inesperadas em derivações como *museu/museleiro*.

O paralelo encontrado entre os resultados dos experimentos e dos dados quantitativos do *corpus* conduz à formulação da hipótese de que a frequência dos padrões fonológicos consiste numa informação relevante para a predição da localização acentual no PB. Isso ocorreria porque, por a frequência de uso impactar e construir as representações fonológicas, os indivíduos utilizariam os padrões fonológicos abstratos presentes na sua gramática a fim de mapear os novos vocábulos por meio de similaridade fonológica. Para isso, eles estenderiam os padrões produtivos arraigados às representações aos novos vocábulos, tornando os tipos silábicos presentes na gramática mais prováveis de emergir do que os tipos silábicos que não apresentam representações fonológicas.

A lacuna fonológica, como será denominada neste trabalho, diz respeito, justamente, à ausência ocasional de uma sequência sonora, com uma certa tonicidade, na gramática fonológica do falante. A lacuna seria ocasional,

uma vez que a sequência sonora respeitaria todas as restrições fonotáticas da língua, não ocorrendo no léxico por mero acaso. Uma vez permitida pela fonotática, a ausência da sequência sonora criaria uma lacuna que diminuiria a probabilidade de sua ocorrência, à medida que a presença de uma sequência sonora similar, com modificação apenas do padrão acentual, ganharia robustez, em decorrência da sua frequência de tipo, facilitando a sua recorrência. A lacuna fonológica diminui, mas não exclui, a possibilidade de que o tipo silábico ausente na gramática seja produzido, posto que processos cognitivos podem interagir preenchendo as lacunas distribucionais.

Pressupõe-se, com isso, que as sequências fonológicas que possuem representações lexicais paroxítonas e oxítonas, [tʃiw], poderiam ser acentuadas em ambas as posições. Em contrapartida, as sequências sonoras que têm apenas uma representação tônica, [niw], tenderiam a ser produzidas com a tonicidade presente no léxico, já que a ausência do tipo silábico com outra tonicidade criaria uma lacuna na gramática fonológica do falante. A maior incidência do acento paroxítono no tipo silábico [tʃiw] em relação a [niw] evidenciaria, portanto, que as representações abstratas dos tipos silábicos abstraídos do uso criam padrões robustos (ou esquemas) que influem na maneira como o acento é produzido.

Tendo em vista que o ataque silábico (ou *onset*) consiste na informação segmental que distingue as duas sequências sonoras e, com isso, o seu comportamento acentual, considera-se, neste estudo, o ataque como uma variável em análise, a qual possui as seguintes variantes: [n] e [tʃ].

Ademais, numa análise aos dados de [tʃiw], percebeu-se que o acento paroxítono só ocorre quando a vogal precedente é oral, especificamente, [a], [ɛ] e [u], como em *versátil*, *projétil* e *inútil*. Quando a vogal é nasal, como em *infantil* e *cantil*, apenas a tonicidade oxítona é encontrada. A distinção do padrão acentual motivada pela vogal que precede [tʃiw] fez com que a informação da qualidade vocálica fosse mais uma informação considerada no grupo de análise 1, com as variantes [V_{oral}] e [V_{nasal}]²⁷.

²⁷ Há uma longa discussão na literatura a respeito do estatuto fonêmico das vogais nasais se monofonêmicas (V), bifonêmicas (V + N), ou mesmo bifásicas (V + N). Por fins de concisão, assumem-se, nesta pesquisa, as vogais nasais como de natureza monofonêmica (V), visto que, segundo Souza & Pacheco (2012, p. 405), "vogais nasais [têm] um comportamento duracional igual ao das vogais orais em sílaba

A discriminação da consoante e da vogal precedente da sequência sonora [iw] motivou a criação dos grupos apresentados em (1) e das pseudopalavras na Tabela 2.7.

- (1) [a.niw], [e.niw] e [u.niw];
 [a.tfiw], [e.tfiw] e [u.tfiw];
 [ẽ.tfiw], [ẽ.tfiw] e [ũ.tfiw].

Tabela 2.7 Pseudopalavras do Grupo de Análise 1.

Grupo de Análise	Pseudopalavra	
[niw]	[bu.da.niw]	[lo.pe.niw]
	[mu.sa.niw]	[ko.xu.niw]
	[si.la.niw]	[pe.tu.niw]
	[ta.ʒe.niw]	[tʃi.ku.niw]
	[li.fe.niw]	
[V _{oral} .tfiw]	[ko.ba.tfiw]	[ve.ne.tfiw]
	[fe.ka.tfiw]	[de.xu.tfiw]
	[si.ma.tfiw]	[na.tu.tfiw]
	[go.le.tfiw]	[tʃi.ʒu.tfiw]
	[su.pe.tfiw]	
[V _{nasal} .tfiw]	[mo.sẽ.tfiw]	[po.lẽ.tfiw]
	[si.pẽ.tfiw]	[ba.pũ.tfiw]
	[ve.nẽ.tfiw]	[bi.kũ.tfiw]
	[ka.bẽ.tfiw]	[xe.bũ.tfiw]
	[ta.vẽ.tfiw]	

Baseadas na hipótese de que a frequência dos tipos silábicos contribui para a determinação da localização acentual, formulam-se, como expõe o Quadro 2.1, as seguintes predições para os tipos silábicos do grupo de análise 1:

O grupo de análise 1 verificará se a frequência de tipo tem impacto nas representações e contribui para a predição da localização acentual. Pressupõe-se, com isso, que os tipos silábicos que possuem mais de uma representação

aberta (VN=V), valores que podem ser utilizados para confirmar a hipótese de que elas sejam monofonêmicas". Para mais detalhes, confira Cagliari (1977), Fonseca (1984), Medeiros (2007) e Souza & Pacheco (2012).

Grupo de Análise 1	[niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw]
[niw]	Somente acento oxítono ocorrerá, posto que todas as palavras terminadas em [niw] no PB são oxítonas. Assim, o único tipo acentual silábico acessível aos falantes do PB para [niw] é o oxítono. Exemplos: <i>canil</i> , <i>sumiu</i> , etc.
[V _{oral} .tʃiw]	Tanto o acento oxítono quanto o paroxítono ocorrerão, uma vez que há palavras no PB terminadas em [V _{oral} .tʃiw] que são oxítonas (ex. <i>sutil</i> , <i>partiu</i>) e paroxítonas (ex. <i>versátil</i> , <i>projétil</i>).
[V _{nasal} .tʃiw]	Somente o acento oxítono ocorrerá, visto que todas as palavras terminadas em [V _{nasal} .tʃiw] no PB são oxítonas (ex. <i>cantil</i> , <i>infantil</i>).

Quadro 2.1 Hipóteses do Grupo de Análise 1.

acentual tenderão a receber as tonicidades existentes em seu léxico, como [V_{oral}.tʃiw], enquanto tipos silábicos que só possuem uma representação acentual, como [niw] e [V_{nasal}.tʃiw], tenderão a receber uma única tonicidade. Se as previsões para o grupo de análise 1 forem confirmadas, elas corroborarão as hipóteses de que os padrões abstratos têm impacto nas representações fonológicas e de que a frequência de tipo e as informações segmentais são relevantes para a determinação do acento em PB. Ainda, mostrarão que a sílaba (com o ataque silábico) consiste numa posição relevante e considerada pelos falantes na atribuição acentual, e não somente o pé métrico (ou a rima), como postulam as propostas métricas.

Grupo de Análise 2: [-or] e [-ar]

A seleção dos tipos silábicos [gor], [tor], [vor], [gar], [tar] e [var]²⁸ para compor o grupo de análise 2 teve como motivação os resultados obtidos no pré-teste²⁹. Nele, verificou-se a emergência de produções paroxítonas em vocábulos terminados em [vor], o qual não possui tipos silábicos paroxítonos

²⁸ Apenas por fins de concisão, o grupo de análise 2 será referido como [-or] e [-ar].

²⁹ Para mais detalhes, verifique o anexo C.

de acordo com o *Corpus* ABG. Os resultados quantitativos encontrados no *corpus* podem ser consultados na Tabela 2.8.

Tabela 2.8 Frequência de tipo das sequências [-or] e [-ar].

Sequência Sonora	Paroxítona	Oxítona
[-or]	10	713
[gor]	1	9
[tor]	5	117
[vor]	0	5
[-ar]	14	2.370
[gar]	0	110
[tar]	0	392
[var]	0	54

Baseada no conceito de lacuna fonológica, a ausência de tipos silábicos [vor] com tonicidade paroxítona, conforme a Tabela 2.8, prediria que o acento paroxítono tenderia a não ser produzido nessa sequência sonora, posto que a falta de uma representação fonológica diminuiria a probabilidade de sua emergência. A manifestação de um padrão não esperado, entretanto, incita questionamentos a respeito das motivações de sua ocorrência, principalmente por ter ocorrido num *corpus* ainda amostral. Trata-se de uma produção inesperada, já que tanto os vocábulos reais terminados em [vor], tais como *fervor*, *pavor*, *favor*, quanto a rima [-or], em muitos casos, sufixo desencadeador de agentividade, como *lutador*, *trabalhador*, *comprador*, ocorrem, majoritariamente, com acento oxítono. A expectativa era que a produtividade do padrão fonológico ['Cor] conduziria à produção do acento oxítono em detrimento do paroxítono, fato que não ocorreu.

Tendo em vista que a língua compreende uso e processos cognitivos, torna-se natural presumir que não só a frequência prediz a localização acentual, mas também que processos cognitivos, tal como a comparação por extensão (ou extensão analógica), possam atuar. Isso ocorreria através da interação entre processos de abstração, que criam os esquemas fonológicos, com processos de comparação, que categorizam ou extensionam os novos padrões aos padrões existentes. O grupo de análise 2 surge, portanto, com o objetivo de investigar

em que medida a produtividade de padrões fonológicos interage com processos analógicos na atribuição de tonicidade através da similaridade lexical.

Os processos analógicos seriam motivados pela frequência de ocorrência de alguns vocábulos que apresentam similaridade fonológica com [vor], como *Vítor* e *Igor*. Segundo Bybee (2010), a analogia torna-se mais provável quanto mais similar forem as conexões fonológicas entre o vocábulo-alvo (as pseudopalavras) e seus vizinhos fonológicos. A frequência de ocorrência do vocábulo-alvo consiste em outro fator que potencializa a analogia, posto que vocábulos de baixa frequência têm conexões mais fracas, sendo mais suscetíveis a sofrerem processos analógicos, do que vocábulos de alta frequência, os quais possuem conexões mais fortes e resistentes a mudanças.

Para que a interação entre produtividade e analogia pudesse ser investigada, determinaram-se para o grupo de análise 2 os tipos silábicos [gor], [tor], [vor] em contraste com [gar], [tar], [var], assumindo a consoante precedente como variável em análise. A escolha de [Car] decorre de sua alta produtividade com acento oxítono, com 2.370 tipos, e, principalmente, da baixa frequência dos padrões fonológicos similares com tonicidade paroxítona, como *ímpar* e *Óscar*³⁰.

Prediz-se, assim, que ambas as sequências sonoras serão produzidas com tonicidade paroxítona e oxítona, entretanto, a incidência de acento paroxítono em [Cor] será significativamente maior do que em [Car]. Isso ocorreria porque a frequência dos padrões fonológicos similares com tonicidade paroxítona de [Cor], como *Vítor* e *Igor*³¹, engatilhará processos de extensão analógica acentual, por similaridade fonológica, entre as palavras reais e as pseudopalavras. Apresentam-se, no Quadro 2.2, as predições para cada tipo silábico em investigação.

Para que as hipóteses elencadas acima fossem testadas, assumiu-se, no grupo de análise 1, a vogal precedente como uma variável relevante. Contudo, para o grupo de análise 2, determinou-se a consoante precedente e a vogal [i] a fim de maximizar a similaridade entre pseudopalavras e vocábulos reais. Diante disso, definiram-se os tipos silábicos apresentados em (2) e as

³⁰ Segundo o *Corpus* ABG, *ímpar* tem frequência de ocorrência correspondente a 6 e *Óscar* a 3.

³¹ Segundo o *Corpus* ABG, a frequência de ocorrência de *Vítor* corresponde a 22 e de *Igor* a 16.

Grupo 2	[-or] e [-ar]
[-or]	Tanto o acento oxítono quanto o paroxítono ocorrerão, uma vez que há palavras no PB terminadas em [-or] que são tanto oxítonas (ex. <i>pavor, fervor</i>) quanto paroxítonas (ex. <i>Vítor, Igor</i>). Haverá maior incidência do acento paroxítono em decorrência da frequência dos vocábulos similares.
[-ar]	Tanto o acento oxítono quanto o paroxítono ocorrerão, posto que há palavras no PB terminadas em [-ar] tanto oxítonas (ex. <i>militar, cantar</i>) quanto paroxítonas (ex. <i>açúcar, Óscar</i>). Entretanto, o acento oxítono será preferido em decorrência da alta frequência de tipo de [-ar].

Quadro 2.2 Hipóteses do grupo de análise 2.

pseudopalavras na Tabela 2.9.

- (2) [i.gor], [i.tor] e [i.vor];
[i.gar], [i.tar] e [i.var];

Tabela 2.9 Pseudopalavras do Grupo de Análise 2.

Grupo de Análise	Pseudopalavra	
[Cor]	[ka.pi.gor]	[zi.bi.tor]
	[me.dʒi.gor]	[de.fi.vor]
	[po.si.gor]	[li.tʃi.vor]
	[ka.mi.tor]	[ta.ni.vor]
	[fe.li.tor]	
[Car]	[ba.zi.gar]	[no.vi.tar]
	[lu.fi.gar]	[lo.ʒi.var]
	[se.tʃi.gar]	[pe.si.var]
	[ga.bi.tar]	[xa.pi.var]
	[du.mi.tar]	

Se as previsões para o grupo de análise 2 forem confirmadas, elas corroborarão a hipótese de que processos analógicos atuam juntamente com a produtividade dos padrões fonológicos a fim de determinar a localização acentual. Além disso, demonstrarão que não só a sílaba final, mas também a vogal que a precede (da sílaba anterior) constituem informações importantes para o acento em português. Dessa maneira, a frequência de tipo, de um lado, seria responsável pela construção das redes de conexões fonológicas. Enquanto a frequência dos padrões fonológicos similares, de outro lado, por meio dos processos de abstração, comparação e esquematização, interagiria na construção de esquemas que extensionam os padrões existentes com a finalidade de preencher as lacunas distribucionais da língua.

2.2.3 Quantidade de Dados

Para cada variante dos grupos de análise 1 e 2, isto é, [niw], [V_{oral}.tʃiw], [V_{nasal}.tʃiw], [Cor] e [Car], construíram-se 9 pseudopalavras, as quais foram repetidas três vezes, de modo aleatório, sendo a mesma sequência de estímulos para todos os informantes. Os grupos 1 e 2 foram testados em simultâneo, totalizando 135 produções por informante (36 indivíduos), 4.860 em conjunto. Pausas foram inseridas ao longo do teste, porém, para evitar possíveis gatilhos acentuais motivados pela condição experimental, não foram utilizadas palavras distratoras.

2.2.4 Procedimentos

De modo similar à condição imagem, apresentada na seção 2.1.3, a tarefa deste experimento consistia na extração de sons por meio de imagens, realizada em duas etapas. Na primeira, fase de reconhecimento de imagens, os participantes eram submetidos à exibição das 73 imagens utilizadas no teste, acompanhadas de sua nomeação por meio de recursos auditivos, isto é, ao mesmo tempo em que o informante visualizava, por exemplo, a imagem de uma *geladeira*, ele ouvia a sequência sonora [ʒe.la'dej.rɐ]. Os indivíduos eram instruídos a ter atenção aos nomes conferidos a cada imagem de modo a assegurar

que, na próxima fase, eles nomeariam as imagens da mesma maneira que a pesquisadora as nomeou.

Na segunda etapa, fase de realização do teste, os sujeitos eram expostos a um rápido treino com três palavras reais, para que eles compreendessem a tarefa a ser realizada. As instruções, também disponibilizadas através de áudio, explicavam e exemplificavam como eles deveriam extrair os sons das palavras que nomeavam as imagens e como comporiam as palavras. Além disso, os participantes eram informados de que as palavras a serem formadas consistiam em palavras não reais, por isso, um estranhamento poderia ocorrer, entretanto, eles deveriam produzir as palavras da forma mais natural possível. Ainda, eles eram informados que, em caso de esquecimento do nome da imagem, a pesquisadora poderia auxiliá-los na nomeação. Ressalta-se que nenhuma ajuda foi conferida à extração dos sons, apenas à nomeação das imagens.

Após o treino e o esclarecimento das dúvidas, eles iniciavam o teste, o qual consistia na apresentação de três imagens sequenciais, a partir das quais os informantes deveriam extrair, de maneira geral, a primeira sílaba das palavras que nomeavam as imagens e, quando perguntados, deveriam produzir a palavra formada. A título de exemplo, [ba] de *bateria*, [zi] de *Ziraldo* e [gar] de *garçom*, sucedida da pergunta auditiva "*Qual palavra foi formada?*", como apresentado na Figura 2.2. Nesse momento, eles deveriam dizer [ba.zi.gar].

Vale a ressalva de que as imagens selecionadas restringiam-se a imagens prototípicas e que os sons extraídos eram átonos (em geral, pré-tônicos), exceto nas pseudopalavras terminadas em [vor] e [var], visto que não foi possível encontrar uma palavra que contivesse essas sequências sonoras e que pudesse ser representada por meio de imagens. Nesses casos, todas as sílabas da pseudopalavra a ser formada eram tônicas. Esse cuidado foi tomado para que a tonicidade da palavra estímulo não fosse gatilho acidental para a pseudopalavra-alvo. Ainda, para minimizar a complexidade da tarefa, procuraram-se imagens que pudessem ter a sua primeira sílaba extraída, porém, em alguns casos, foi necessária a utilização da segunda ou da última

Figura 2.2 Exemplo do experimento com a pseudopalavra [ba.zi.gar].

sequência sonora³².

O experimento foi realizado em sala comum com a presença apenas da pesquisadora e do informante, com duração média de 30 minutos. Todas as produções auditivas foram gravadas no programa Audacity (Audacity Team, 2016), a uma taxa de amostragem de 44.100 Hz e a 16 bits por amostra. As imagens foram exibidas por intermédio do programa Microsoft Power Point, com uma imagem por *slide*, no Notebook HP Compact. O fone de ouvido utilizado consistiu no Sennheiser, modelo HD201. As instruções e o nome das imagens foram gravados no estúdio da Faculdade de Letras da Universidade de São Paulo, produzidos por um falante nativo do português brasileiro, nascido na região Metropolitana de São Paulo, com boa dicção e sem patologias relacionadas à fala.

2.3 Método de Análise dos Dados

2.3.1 Análise Perceptiva

Todos os áudios foram submetidos a três audições pela pesquisadora e foram codificados de acordo com a tonicidade - proparoxítono, paroxítono ou oxítono - de cada pseudopalavra. Somente as pseudopalavras que apresentaram

³² São elas: [kũ], [pũ], [bũ], [nẽ], [var], [vor] e [tʃiw].

divergência em qualquer uma das oitavas foram submetidas à análise acústica. Nela, mediu-se, por meio do *software* Praat (BOERSMA & WEENINK, 2013), o principal correlato acústico do acento primário no português: a duração, para definir a localização do acento primário (cf. MARTINS, 1988; MASSINI-CAGLIARI, 1992; CONSONI, 2010).

2.3.2 Variáveis e Modelos Estatísticos

Com a finalidade de inquirir se os tipos silábicos alvos deste estudo são relevantes para a predição da localização acentual no PB, selecionaram-se, como variável dependente, a *tonicidade*³³, e como variáveis independentes, o *grupo de análise*, a *sílabas final*, a *repetição* e a *pseudopalavra*, e, como variável aleatória, o *informante*.

A variável dependente *tonicidade* codifica o resultado do experimento realizado, isto é, por meio de suas variantes proparoxítone, paroxítone e oxítone, ela marca o padrão acentual empregado pelo falante no momento da produção dos estímulos. As possíveis motivações definidas neste estudo para a alternância do padrão acentual consistem nas variáveis independentes. A mais relevante e a central para esta investigação consiste em *grupo de análise*, já que ela compreende os tipos silábicos. Ela é central, posto que a hipótese deste estudo será comprovada ou refutada com a sua seleção e significância nos testes estatísticos. Assim, as hipóteses gerais, para os grupos de análise 1 e 2, podem ser definidas como:

H^1 = O *grupo de análise* determina a localização acentual.

H^0 = O *grupo de análise* NÃO determina a localização acentual.

Caso o valor de p, ou de significância, do *grupo de análise* seja inferior a 0,05 - valor comumente utilizado nas Ciências Humanas -, a hipótese alternativa (H^1) será selecionada, demonstrando assim que os tipos silábicos têm papel relevante na predição do acento em PB. Em contrapartida, se o valor de p for superior a 0,05, a hipótese nula (H^0) não poderá ser descartada, fato que trará resultados inconclusivos a respeito do papel da variação em

³³ Por fins de clareza, todas as variáveis serão referidas em *italico*.

relação ao acento em PB. As variantes de *grupo de análise*, no grupo de análise 1, são [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw] e, no grupo de análise 2, são [gor], [tor], [vor], [gar], [tar] e [var].

Além do *grupo de análise*, determinaram-se outras variáveis que podem influir no comportamento da *tonicidade*, são elas: *sílabo final*, *repetição* e *pseudopalavra*. A *sílabo final*, no grupo de análise 1, codifica a informação da vogal que precede o tipo silábico final e, no grupo de análise 2, codifica a consoante e a vogal que precedem a rima final. A variável *repetição* refere-se a qual produção, se 1^a, 2^a ou 3^a, a pseudopalavra corresponde no experimento. Ela foi inserida para que a constância do comportamento dos falantes pudesse ser observada e, também, para analisar possíveis interferências experimentais. A última variável independente consiste nos próprios estímulos do experimento, as *pseudopalavras*. Embora elas tendam a ser significativas em decorrência da sua variabilidade, já que são 45 estímulos diferentes, uma análise por *pseudopalavra* permite visualizar se um determinado vocábulo favorece algum padrão acentual.

Ademais, os *informantes* consistem em outra variável importante para os estudos experimentais. De modo geral, eles são caracterizados como variáveis aleatórias, posto que não podem ser controlados ou manipulados pelos pesquisadores. No entanto, eles são essenciais para garantir adequação e precisão aos modelos estatísticos, como ocorreu neste estudo.

Todas as variáveis dependente e independentes, descritas acima, com suas propriedades e variantes nos grupos de análise 1 e 2 são sintetizadas e expostas nos Quadros 2.3 e 2.4, respectivamente.

A fim de investigar o papel que cada variável independente, apresentada nos Quadros 2.3 e 2.4, exerce na predição da variável dependente *tonicidade*, análises monofatoriais com o teste qui-quadrado (χ^2) foram realizadas. Além disso, para que as variáveis pudessem ser testadas em interação umas com as outras, realizaram-se testes de regressão logística binomial, multifatorial e com efeitos aleatórios. Todos os testes estatísticos foram executados no *software* R³⁴ (R Core Team, 2015).

³⁴ Todos os testes foram realizados com o auxílio do estatístico Helder Rojas.

2. METODOLOGIA

	Variável	Tipo	Variantes
Tonicidade	Dependente	Nominal	Proparoxítona Paroxítona Oxítona
Grupo de Análise	Independente	Nominal	[niw] [V _{oral} .tʃiw] [V _{nasal} .tʃiw]
Sílaba Final	Independente	Nominal	[a.niw], [e.niw], [u.niw] [a.tʃiw], [e.tʃiw], [u.tʃiw] [ẽ.tʃiw], [ẽ̃.tʃiw], [ũ.tʃiw]
Repetição	Independente	Nominal	1 ^a 2 ^a 3 ^a
Pseudopalavra	Independente	Nominal	[bu.da.niw], [ta.ʒe.niw], ... [ko.ba.tʃiw], [go.le.tʃiw], ... [mo.sẽ.tʃiw], [ka.bẽ.tʃiw], ...
Informante	Aleatória	Nominal	F1, F2, F3, ..., F36

Quadro 2.3 Variáveis dependente e independentes com suas propriedades e variantes do Grupo de Análise 1.

2.4 Considerações Finais

Este capítulo apresentou a metodologia empregada na elaboração do experimento deste estudo, explicitando as motivações e as justificativas para a adoção de cada procedimento metodológico. Mostraram-se, assim, os desafios metodológicos que conduziram a realização de um metaexperimento e a compilação de um *corpus*, bem como os diferentes perfis dos 36 informantes que participaram desta pesquisa, a forma como os estímulos foram compostos e as hipóteses que abrangeram a delimitação dos tipos silábicos. Além disso, apresentaram-se os procedimentos realizados para a elicitación e codificação dos dados, as variáveis, as hipóteses e os modelos estatísticos utilizados na análise dos dados.

	Variável	Tipo	Variantes
Tonicidade	Dependente	Nominal	Proparoxítona Paroxítona Oxítona
Grupo de Análise	Independente	Nominal	[Cor] [Car]
Sílaba Final	Independente	Nominal	[i.vor], [i.gor], [i.tor] [i.var], [i.gar], [i.tar]
Repetição	Independente	Nominal	1 ^a 2 ^a 3 ^a
Pseudopalavra	Independente	Nominal	[ta.ni.vor], [po.si.gor], ... [lo.ʒi.var], [lu.fi.gar], ...
Informante	Aleatória	Nominal	F1, F2, F3, ..., F36

Quadro 2.4 Variáveis dependente e independentes com suas propriedades e variantes do Grupo de Análise 2.

3

Resultados e Discussão

Este capítulo apresenta e discute os resultados obtidos no experimento descrito no capítulo anterior, o qual contou com dois grupos de análise: i. [niw], [V_{oral}.tʃiw], [V_{nasal}.tʃiw]; e ii. [-or], [-ar]. O objetivo principal ao desenvolver o procedimento experimental foi o de investigar o papel da frequência dos tipos silábicos na atribuição acentual em pseudopalavras no português brasileiro. Este capítulo apresenta e discute os resultados obtidos à luz das hipóteses de atribuição do acento apresentadas no Capítulo 2.

A análise experimental, que se restringiu a sequências segmentais que formariam uma palavra do tipo CV-CV-CVC, dividiu-se em dois grupos de análise. O grupo de análise 1, composto pelos tipos silábicos [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], investigou em que medida a frequência dos tipos silábicos pode prever a localização acentual. O grupo de análise 2, composto pelos tipos silábicos [-or] e [-ar], inquiriu se a frequência dos tipos silábicos interagiria com efeitos de similaridade lexical em decorrência da frequência dos vizinhos fonológicos.

Para que as hipóteses deste estudo fossem testadas, determinaram-se como variável dependente a *tonicidade*¹, como variáveis independentes o *grupo de análise*, a *sílabas final*, a *pseudopalavra* e a *repetição*, e como variável aleatória o *informante*. A relevância de cada variável foi testada isoladamente, em análises monofatoriais, e em interação com as demais variáveis, em análises multifatoriais. Os modelos empregados nas análises consistiram em testes de

¹ Por fins de clareza, as variáveis serão referidas em *itálico*.

qui-quadrado (χ^2), Modelo de Regressão Logística Binomial, Multinomial e com Efeitos Mistos, os quais foram processados no *software* R (R Core Team, 2015).

A seção 3.1 apresenta os resultados gerais desta pesquisa. As seções 3.2 e 3.3 apresentam e discutem os resultados obtidos, respectivamente, nos grupos de análise 1 e 2. As seções 3.4 e 3.5 dedicam-se à discussão qualitativa das produções proparoxítonas e da variável *informante*, respectivamente. Por fim, na seção 3.6, expõem-se os resultados gerais obtidos com a realização deste experimento.

3.1 Resultados Gerais

Todas as produções dos falantes apresentaram correlatos fonéticos do acento primário do PB, com uma maior duração e intensidade de produção das vogais tônicas (cf. MARTINS, 1988; MASSINI-CAGLIARI, 1992; CONSONI, 2010). Embora nesta pesquisa o caráter fonético do acento tenha sido avaliado de oitiva, ou seja, sem interpretação experimental, ele é relevante ao indicar que falantes expressam o correlato acentual de sua língua materna ao produzirem pseudopalavras. Pesquisas futuras poderão avaliar se quando os indivíduos são apresentados a dados de uma língua estrangeira que tenha, por exemplo, vogais curtas e longas, o falante do PB empregará as mesmas propriedades fonéticas acentuais utilizadas em sua língua materna. A expectativa é que este seja o caso. O desafio será investigar a evolução da proficiência de falantes quanto aos correlatos fonéticos do acento e à aquisição dos aspectos segmentais particulares de L2, como as vogais longas e curtas em posição acentuada.

Quanto à atribuição do acento primário nas pseudopalavras, destacaram-se, no Capítulo 1, quatro tendências para a atribuição acentual, como indicado no Quadro 3.1.

Hipótese	Atribuição de acento primário
Lexical	Ao acaso
Métrica e Morfológica	Oxítono
Multirrepresentacional	Qualquer padrão acentual da língua, sob efeitos de frequência

Quadro 3.1 Predições de atribuição acentual por hipótese de análise.

A partir do Quadro 3.1, pode-se sugerir que haverá uma tendência ao acento oxítono nos resultados. Essa tendência ocorreria visto que, dada a estrutura CV-CV-CVC, a hipótese Métrica prevê que, em decorrência da sensibilidade ao peso silábico, apenas o acento oxítono seria produzido. De modo semelhante, para a hipótese Morfológica, a ausência da vogal temática faria com que o acento incidisse na última vogal do radical, com acento oxítono. Já a hipótese Multirrepresentacional prediz que possam ocorrer oxítonos, concomitante com o acento proparoxítono e paroxítono. Por fim, a hipótese Lexical, ao assumir que o acento é atribuído ao acaso, prediz que qualquer padrão acentual possa emergir. Considere os dados da Tabela 3.1.

Tabela 3.1 Atribuição acentual nos dados gerais.

Proparoxítona	Paroxítona	Oxítona
53 (1,11%)	1.175 (24,54%)	3.560 (74,35%)

$$\chi^2 = 4019.7 (2), p < 0.0001$$

Na Tabela 3.1, observa-se 1,11% de produções proparoxítonas, 24,54% de paroxítonas e 74,35% de oxítonas, com $\chi^2 = 4019.7 (2)$ e $p < 0.0001$. Estes resultados indicam o favorecimento do acento oxítono em detrimento do acento proparoxítono ou paroxítono. O acento proparoxítono ocorre com baixo índice, sendo registrado em 1,11% dos dados. Este resultado pode indicar que quando expostos a pseudopalavras os falantes utilizam qualquer um dos padrões acentuais possíveis do PB: proparoxítono, paroxítono ou oxítono. Contudo, essa atribuição não se dá ao acaso nem de forma categórica, na

medida em que há um favorecimento do acento oxítono (74,35%) em relação aos demais padrões acentuais, o que oferece desafios para as hipóteses Métrica e Morfológica, as quais predizem exclusivamente o acento oxítono, ou seja, a atribuição acentual proparoxítona ou paroxítona não seria esperada. Assim, um importante resultado expresso a partir dos dados da Tabela 3.1 é que qualquer um dos três padrões acentuais do PB - proparoxítono, paroxítono ou oxítono – pode ocorrer. O desafio é buscar entender as condições em que o acento proparoxítono e paroxítono ocorrem, uma vez que o favorecimento de oxítonas já seria esperado.

Ao considerar o baixo índice de produções com padrão acentual proparoxítono, a análise e a discussão apresentadas neste capítulo consideram apenas os parâmetros para a atribuição do acento oxítono e paroxítono. Às proparoxítonas é reservada uma discussão na seção 3.4.

3.2 Grupo de Análise 1: [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw]

A presente seção expõe e discute os resultados obtidos no grupo de análise 1, o qual compreende as variantes [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw]. Esse grupo investigou, por meio do conceito de lacuna fonológica, em que medida a frequência dos tipos silábicos pode prever a localização acentual. Pautado nos Modelos Multirrepresentacionais, pressupôs-se que o uso, por ser responsável pela construção dos esquemas fonológicos, pode impactar na maneira como o acento é produzido em PB. Dessa forma, os tipos silábicos que apresentam mais de uma representação fonológica tendem a ter uma maior quantidade de produções com as tonicidades presentes em sua gramática do que os tipos silábicos que só possuem uma representação.

Esperava-se, assim, que a sequência sonora [niw], por apresentar somente representações fonológicas oxítonas, 19 tipos², teria maior incidência de acento oxítono em relação ao acento paroxítono, como em [ʒu.ve'niw] em detrimento

² Verifique a frequência de tipo dessas sequências sonoras na Tabela 3.2 a partir de dados do *Corpus ABG*.

de *[zu've.niw]. Seguindo a mesma predição, a sequência [V_{nasal}.tʃiw], por possuir 12 tipos oxítonos e 0 paroxítonos, também teria predominância de tonicidade final, como [kẽ'tʃiw] em detrimento de *['kẽ.tʃiw]. Em contrapartida, [V_{oral}.tʃiw] manifestaria-se com ambas as tonicidades, já que possui tipos com os dois padrões acentuais, 46 oxítonas e 14 paroxítonas, como [ĩ.fẽ'tʃiw] e [ver'sa.tʃiw].

As hipóteses estabelecidas para o grupo de análise 1 são retomadas no Quadro 3.2. Vale a ressalva de que apesar de as hipóteses terem que ser categóricas para a estatística, os Modelos Multirrepresentacionais não preveem resultados precisos assim. Ao contrário, espera-se que eles reflitam os padrões estocásticos da língua inferidos através dos índices de frequência, os quais podem interagir com processos cognitivos. Basicamente, as hipóteses que seguem são base apenas para direcionar os comportamentos acentuais por meio da frequência.

Grupo de Análise 1	[niw], [V _{oral} .tʃiw] e [V _{nasal} .tʃiw]
[niw]	Somente acento oxítono ocorrerá, posto que todas as palavras terminadas em [niw] no PB são oxítonas. Assim, o único tipo acentual silábico acessível aos falantes do PB para [niw] é o oxítono. Exemplos: <i>canil</i> , <i>sumiu</i> , etc.
[V _{oral} .tʃiw]	Tanto o acento oxítono quanto o paroxítono ocorrerão, uma vez que há palavras no PB terminadas em [V _{oral} .tʃiw] que são oxítonas (ex. <i>sutil</i> , <i>partiu</i>) e paroxítonas (ex. <i>versátil</i> , <i>projétíl</i>).
[V _{nasal} .tʃiw]	Somente o acento oxítono ocorrerá, visto que todas as palavras terminadas em [V _{nasal} .tʃiw] no PB são oxítonas (ex. <i>cantil</i> , <i>infantil</i>).

Quadro 3.2 Hipóteses dos grupos de análise 1.

Considere a Tabela 3.2, a qual apresenta, do lado esquerdo, a frequência de cada tipo silábico com acento paroxítono e oxítono. Os valores correspondem à quantidade de tipos de palavra presentes no *Corpus* ABG. Pautando-se

no conceito de lacuna fonológica, os dados sustentam a predição de que, se o uso impacta nas representações, serão encontradas somente produções oxítonas em [niw] e [V_{nasal}.tʃiw] e paroxítonas e oxítonas em [V_{oral}.tʃiw]. Como apresentado na Tabela 3.2, os resultados obtidos não acolhem tal predição, como será discutido adiante.

Tabela 3.2 Frequência de tipo das sequências [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw] por *tonicidade* no *Corpus* ABG e nos resultados do Grupo de Análise 1.

Sequência Sonora	<i>Corpus</i> ABG		Resultados Obtidos ³	
	Paroxítona	Oxítona	Paroxítona	Oxítona
[niw]	0 (0%)	19 (100%)	182 (19%)	777 (80,9%)
[V _{oral} .tʃiw]	14 (23,3%)	46 (76,7%)	309 (32,2%)	651 (67,7%)
[V _{nasal} .tʃiw]	0 (0%)	12 (100%)	332 (34,7%)	625 (65,3%)

$$\chi^2 = 67.889 (4) \text{ e } p < 0.0001$$

Nos dados gerais, verificou-se, conforme demonstra a Tabela 3.3, que o acento oxítono ocorreu em maior proporção, 70,4% dos casos, seguido pelas paroxítonas com 28,22% e pelas proparoxítonas com 0,07%. A diferença mostra-se significativa com $\chi^2 = 5474.1 (4)$ e $p < 0.0001$.

Tabela 3.3 Atribuição acentual no grupo de análise 1 - [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw].

Tonicidade	Ocorrências (%)
Proparoxítona	2 (0,07%)
Paroxítona	823 (28,22%)
Oxítona	2.053 (70,4%)
Excluídos	38 (1,3%)

$$\chi^2 = 5474.1 (4) \text{ e } p < 0.0001$$

Os resultados obtidos no grupo de análise 1 seguem a tendência apontada pelas propostas fonológicas de que o acento oxítono seria predominante em palavras terminadas com sílaba final pesada, fato que ocorreu em 70,4% dos casos. Contudo, o que parte das propostas não previu foram os, aproximadamente, 28% de produções paroxítonas e os 0,07% de proparoxítonas.

³ Em [niw] e [V_{oral}.tʃiw], os dados não somam 100%, visto que foi encontrada uma produção proparoxítona para cada uma delas, o que se refere a 0,1% dos dados.

De acordo com a hipótese Métrica, a sílaba final pesada [iw] das pseudopalavras conduziria a produções exclusivamente oxítonas. Por mais que a hipótese Métrica postule a extrametricidade segmental ou silábica para justificar as produções divergentes da esperada (oxítona), tal mecanismo só poderia ser aplicado a palavras reais. As pseudopalavras não poderiam ser abarcadas na extrametricidade, posto que elas, por não serem palavras reais, não possuem uma representação lexical, isto é, uma marcação prévia de invisibilidade segmental ou silábica, como requerido em <niw> e <tʃiw>. Além disso, não é possível prever que a extrametricidade, no caso das pseudopalavras, ocorreria ao acaso, na medida em que, como será apresentado adiante, produções com atribuição de acento paroxítono seguem uma tendência específica.

Para a hipótese Morfológica, a ausência da vogal temática no final dos vocábulos faria com que o acento incidisse na última vogal do radical, sendo assim, todas as pseudopalavras deste estudo deveriam receber acento oxítono, já que elas terminam em ditongo [iw]. Às produções inesperadas, proparoxítonas e paroxítonas, Lee (1995) postula uma regra que atingiria as palavras marcadas. Note que, de modo semelhante à hipótese Métrica, o conceito de marcação⁴ requer um conhecimento prévio do vocábulo, caso que, novamente, não é aplicável a pseudopalavras. Além disso, tanto a hipótese Morfológica quanto a hipótese Métrica apresentam limitações para justificar como diferentes tonicidades podem ser atribuídas a um mesmo vocábulo, por um mesmo falante, como em [ve'ne.tʃiw] e [ve.ne'tʃiw].

Com relação à hipótese Lexical, ao prever que o acento é livre, Câmara Jr. (1970 [2001]) prevê uma atribuição acentual aleatória, sem qualquer tendência. A aleatoriedade das atribuições acentuais, proposta pelo autor, não é verificada neste experimento, visto que os 70,4% de produções oxítonas e os 28,22% de paroxítonas, as quais, como será exposto adiante, expressam tendências de frequência, questionam a hipótese de aleatoriedade. Ademais, segundo os pressupostos da hipótese Lexical, os falantes não poderiam atribuir acento em pseudopalavras, já que o acento desses vocábulos não teria sido aprendido *a priori*.

⁴ Para mais detalhes, retorne à seção 1.3.3, do Capítulo 1.

Quanto à hipótese Multirrepresentacional, ao propor que generalizações são realizadas a partir de experiências de uso e que elas constroem as representações que expressam os padrões abstratos, os quais podem ser estendidos a novos vocábulos, a proposta de Cantoni (2013) possibilita que todos os padrões acentuais (proparoxítono, paroxítono e oxítono) possam ocorrer, os quais refletem as tendências de frequência da língua.

As hipóteses deste grupo de análise 1, ao se valerem da predição de que os padrões fonológicos abstratos seriam responsáveis pela emergência e extensão do acento a pseudopalavras, permitiam que houvesse produções variáveis, sobretudo se os, aproximadamente, 28% de acento paroxítono compreendessem o tipo $[V_{oral}.t.fiw]$, posto que são encontrados no léxico do PB tanto vocábulos com acento paroxítono quanto com oxítono. Entretanto, observou-se variabilidade acentual, conforme exibido na Tabela 3.2, em todos os tipos silábicos investigados: $[niw]$, $[V_{oral}.t.fiw]$ e $[V_{nasal}.t.fiw]$. A emergência de padrões acentuais inesperados traz indícios de que, embora com índices mais baixos, as lacunas distribucionais presentes na gramática fonológica dos falantes tendem a ser preenchidas em decorrência de processos extensionais entre os padrões abstratos e os tipos silábicos, os quais são realizados através de processos cognitivos de comparação, abstração e extensão⁵.

⁵ Os processos extensionais não devem ser compreendidos como uma regra que engloba irregularidades, visto que a emergência de um padrão, mesmo que inesperado, reflete as possibilidades fonológicas da língua, a qual se vale da produtividade de certos padrões abstratos - no caso, $[-iw]$ - e de processos cognitivos - comparação, abstração e extensão - a fim de preencher as lacunas existentes.

Figura 3.1 Esquema de conexões fonológicas entre [niw] e [tʃiw].

De acordo com os Modelos Multirrepresentacionais, as lacunas distribucionais, conforme demonstra a Figura 3.1, podem ser preenchidas por meio de redes ou de esquemas que emergem a partir de experiências de uso, os quais, através de processos de abstração e de similaridade fonológica, se conectam e se fortalecem, conduzindo ao arraigamento das estruturas fonológicas. Diferentes padrões fonológicos locais, tais como ['niw] e ['tʃiw], representados por círculos com linhas contínuas azuis, ao serem abstraídos geram padrões abstratos, como ['Ciw] e ['σ.Ciw], representados por círculos com linhas contínuas pretas. Os esquemas tornam-se complexos na medida em que os padrões abstratos, os quais, em geral, são produtivos, competem com os padrões locais, os quais apresentam informações fonológicas mais ricas e finas. Os padrões produtivos, por um lado, atraem os vocábulos de baixa frequência; os padrões locais, por outro, potencializados pela semelhança fonológica, conduzem à emergência de subpadrões e de processos extensionais. Todos os padrões abstraídos correspondem a um esquema cognitivo arraigado na gramática fonológica do falante, o qual associa uma sequência sonora, [Ciw], a um

padrão acentual. A conexão será mais forte, arraigada e proeminente quanto mais alta for a sua frequência de tipo.

Nessa perspectiva de análise, o predomínio do acento oxítono no grupo de análise 1 seria resultante de efeitos de frequência de tipo, já que a sequência sonora [iw] com tonicidade oxítona possui 334 tipos silábicos, enquanto a mesma sequência com acento paroxítono possui apenas 50 tipos. A diferença de frequência de tipo é representada na Figura 3.1 pela espessura das nuvens de exemplares. Nuvens mais grossas significam sequências sonoras mais frequentes, nuvens mais finas, menos frequentes. A robustez da conexão fonológica, ou do padrão abstraído, faria com que [iw] se tornasse gatilho analógico das pseudopalavras. Ou seja, o padrão acentual oxítono seria estendido do tipo frequente para as pseudopalavras, as quais não têm frequência na língua. Todavia, essa extensão não seria categórica em decorrência do modo como a frequência de tipo é composta, já que a emergência de um padrão produtivo como [iw] decorre da união de diferentes subpadrões, ou padrões locais, como ['niw], ['tʃiw], ['σ.Ciw], o que poderia resultar tanto em produções oxítonas quanto paroxítonas, como ocorreu.

No caso de [V_{oral}.tʃiw], a presença de dois tipos silábicos, paroxítono e oxítono, na gramática fonológica do falante, possibilitaria com que ambos os padrões acentuais emergissem. No entanto, as produções paroxítonas não esperadas em [niw] e [V_{nasal}.tʃiw] trazem indícios de que há outros fatores que determinam a localização acentual que vão além da sensibilidade ao peso silábico, como prediz a proposta métrica, e da ausência de marcador de palavra, como afirma a proposta morfológica. O objetivo deste experimento consistiu justamente em avaliar se a gramática fonológica, consolidada pelo uso, atua na maneira como o acento é produzido pelos falantes.

Com a finalidade de verificar se o uso e os efeitos de frequência impactam na atribuição do acento, as análises monofatoriais e multifatoriais realizadas consideraram apenas as ocorrências de acento paroxítono e oxítono, visto que houve somente 2 ocorrências proparoxítonas, as quais serão discutidas em detalhes na seção 3.4. A reestruturação dos dados resultou em 28,62% de paroxítonas e 71,38% de oxítonas, preservando a significância com $\chi^2 = 526.04$ (1) e $p < 0.0001$, como demonstra a Tabela 3.4.

Tabela 3.4 Resultados reestruturados de atribuição acentual em [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw] (sem proparoxítonas).

Tonicidade	Ocorrências (%)
Paroxítona	823 (28,62%)
Oxítona	2.053 (71,38%)
$\chi^2 = 526.04$ (1) e $p < 0.0001$	

As análises monofatoriais, realizadas com testes de qui-quadrado (χ^2), investigaram o papel que cada variável independente desempenhava na predição da variável dependente (*tonicidade*). As variáveis independentes analisadas foram: *grupo de análise*, o qual consiste nos tipos alvo desta investigação [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw]; *sílabas final*, que compreende a informação da vogal que precede o tipo silábico final; *pseudopalavra*, a qual consiste nas pseudopalavras alvo desta pesquisa; e *repetição*, que informa se a pseudopalavra produzida corresponde a 1ª, 2ª ou 3ª ocorrência. Os resultados apontaram que todas as variáveis independentes são relevantes para a explicação da variável dependente, *tonicidade*. O valor obtido em cada variável independente no teste de χ^2 com seu respectivo valor de significância pode ser visualizado na Tabela 3.5.

Tabela 3.5 Correlação entre a variável dependente e as variáveis independentes em [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw].

Variáveis	χ^2 (df)	Valor de p
<i>Grupo de Análise</i>	66.897 (2)	< 0.0001
<i>Sílabas Final</i>	74.935 (8)	< 0.0001
<i>Pseudopalavra</i>	92.864 (26)	< 0.0001
<i>Repetição</i>	8.1789 (2)	0.01675

As análises multifatoriais, realizadas com modelo de regressão logística binomial, verificaram quais variáveis independentes atuavam conjuntamente na explicação da variável dependente. Para que as análises fossem realizadas com adequação, as variáveis independentes deveriam ser, realmente, independentes umas das outras, isto é, a informação de uma variável não poderia estar contida na outra. Como as informações de *grupo de análise* e *sílabas final*

estão abarcadas em *pseudopalavra*, essas variáveis tiveram que ser separadas em dois grupos. O primeiro grupo incluiu *grupo de análise* e *sílaba final*, já o segundo grupo, *pseudopalavra*.

Apesar da separação em grupos distintos, quer o primeiro modelo, [*grupo de análise* + *sílaba final* + *repetição*], quer o segundo, [*pseudopalavra* + *repetição*], não obtiveram valores de adequação satisfatórios no modelo de regressão logística, conforme demonstra a Figura 3.2⁶. A falta de adequação é visualizada na comparação entre os erros do modelo (círculos pretos) e a curva de normalidade da distribuição (linha pontilhada). A distância entre eles demonstra que os resultados estimados pelo modelo não correspondem aos resultados obtidos no experimento, o que evidencia a baixa precisão do modelo na explicação dos dados observados.

Figura 3.2 Adequação do modelo de regressão logística em [niw], [V_{oral}.t.fiw] e [V_{nasal}.t.fiw].

Ao buscar a causa da disparidade entre os valores observados e estimados, verificou-se, numa análise de erros por informante, que eles formavam *clusters* hierárquicos com diferentes comportamentos acentuais. De modo geral, observou-se que os informantes apresentavam três comportamentos distintos: i. produziam (majoritariamente) acento oxítono; ii. produziam (majoritariamente) acento paroxítono; e iii. produziam (igualmente) acento paroxítono e oxítono. Essas tendências podem refletir problemas com o desenho experimental, que poderá ser revisto em experimentos futuros.

⁶ Veja o gráfico completo do modelo de regressão logística de [niw], [V_{oral}.t.fiw] e [V_{nasal}.t.fiw] no anexo D.

A diversidade de comportamento acentual conduziu à inserção dos informantes como variável aleatória, já que eles não podem ser controlados pela pesquisadora. O modelo utilizado consistiu no modelo de regressão logística binomial com efeitos aleatórios, o qual obteve uma boa adequação e normalidade, conforme a Figura 3.3⁷, selecionando as variáveis [*grupo de análise + sílaba final + repetição*] e [*pseudopalavra + repetição*] como variáveis relevantes para a atribuição acentual. O resultado completo dos modelos pode ser verificado no anexo F.

Figura 3.3 Erros do modelo de regressão logística binomial com efeitos mistos em [niw], [V_{oral}.tfiw] e [V_{nasal}.tfiw].

O primeiro modelo, que testou [*grupo de análise + sílaba final + repetição + informante*], selecionou todas as variáveis independentes como significativas para a *tonicidade*. Para *grupo de análise*, constatou-se que [niw] favorece o acento oxítono, enquanto, [V_{oral}.tfiw] e [V_{nasal}.tfiw], o acento paroxítono, com $p < 0.0001$, conforme demonstra a Tabela 3.6. Verificou-se, ainda, que há diferença significativa entre [niw], [V_{oral}.tfiw] e [V_{nasal}.tfiw], com $\chi^2 = 66.897(2)$ e $p < 0.0001$, mas não entre [V_{oral}.tfiw] e [V_{nasal}.tfiw], com $\chi^2 = 1.2402(1)$ e $p = 0.2654$.

A significância de *grupo de análise* conduz à seleção da hipótese alternativa (H^1), a qual prediz que a frequência dos tipos silábicos contribui para a localização acentual em pseudopalavras. Essa postulação pôde ser

⁷ Veja o gráfico completo do modelo de regressão logística binomial com efeitos aleatórios de [niw], [V_{oral}.tfiw] e [V_{nasal}.tfiw] no anexo E.

⁸ Observe que as estimativas com valores positivos, êxito, significam preferência oxítona, enquanto os valores negativos, fracasso, predileção paroxítona.

Tabela 3.6 Resultados do modelo de regressão logística com efeitos mistos em [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw].

	Estimativa⁸	Erro Padrão	t	p
[niw]	2.12985	0.30915	6.889	6.88e-12***
[tʃiw]	-1.78543	0.33174	-5.382	7.97e-08***
[n.tʃiw]	-2.43928	0.31838	-7.662	2.50e-14***

verificada com o favorecimento de [niw] ao acento oxítono e de [V_{oral}.tʃiw] ao paroxítono, resultados que oferecem evidências a favor da hipótese de que o uso e a frequência dos tipos silábicos têm impacto nas representações fonológicas, apesar dos dados divergentes de [V_{nasal}.tʃiw]⁹. Resultados semelhantes já foram atestados em outros estudos. Vitevitch et al (1997), por exemplo, constataram que a frequência dos padrões fonológicos é relevante no julgamento de gramaticalidade e no reconhecimento dos vocábulos, posto que vocábulos que possuíam frequências fonotáticas mais altas eram julgados e processados mais rapidamente do que vocábulos que possuíam frequências fonotáticas com menores índices.

Os resultados obtidos no grupo de análise 1 seguem, parcialmente, as hipóteses iniciais deste estudo. Para [niw], observou-se, como esperado, preferência pelo acento oxítono com algumas produções paroxítonas, as quais podem refletir extensões de padrões como *hábil* e *portátil*. Com relação a [V_{oral}.tʃiw] e a [V_{nasal}.tʃiw], o favorecimento do acento paroxítono era esperado somente no tipo [V_{oral}.tʃiw], visto que ele apresenta representações tanto paroxítonas quanto oxítonas no léxico da língua, fator que motivaria a sua preferência. Esperava-se, todavia, que a ausência de tipos [V_{nasal}.tʃiw] paroxítonos reduziria a probabilidade de sua ocorrência, como constatado em [niw], fato que não ocorreu. A predileção inesperada pelo acento paroxítono de [V_{nasal}.tʃiw] pode ter resultado da ausência de distinção da vogal precedente entre os tipos [V_{oral}.tʃiw] e [V_{nasal}.tʃiw]. Somado a isso, vocábulos com tipos similares, como *portátil* e *hábil*, podem ter estendido os seus padrões acentuais aos novos vocábulos.

⁹ Adiante, argumenta-se para a hipótese de que os falantes interpretaram [V_{oral}.tʃiw] e [V_{nasal}.tʃiw] da mesma maneira.

A similaridade no comportamento da $[V_{nasal}.tʃiw]$ com $[V_{oral}.tʃiw]$ indicaria que os falantes, possivelmente, não fazem distinção da qualidade da vogal que precede o tipo silábico final. Uma das razões para esse comportamento pode estar relacionada à forma como os estímulos foram processados. Para Vitevitch et al (1999), diferentemente das palavras reais, as quais são processadas no nível lexical, já que estão presentes nas representações fonológicas, as pseudopalavras são processadas em níveis menores do que a palavra, em nível sublexical, com diferentes n-gramas (bigramas, trigramas, etc).

A propensão a processamentos sublexicais, a princípio, possibilitaria que os falantes fizessem associações em diferentes níveis de abstração (pé, sílaba, rima). No entanto, o experimento deste estudo, ao solicitar a extração do primeiro som da palavra que nomeava as imagens (i. é, da primeira sílaba), de certo modo, condicionava os falantes a processarem as unidades no nível silábico, restringindo a análise de níveis maiores ou menores do que a sílaba¹⁰. A tarefa de extração dos sons, dessa forma, direcionava o falante a fazer abstrações, extensões e categorizações apenas no nível silábico. O padrão abstraído, $[tʃiw]$, ao associar uma sequência sonora a um padrão acentual, paroxítono e oxítono, reflete os padrões presentes na gramática do falante, não permitindo, com isso, que os indivíduos realizassem abstrações que considerassem a vogal precedente, como este estudo esperava. Diante disso, a semelhança no comportamento de $[V_{oral}.tʃiw]$ e $[V_{nasal}.tʃiw]$ seria coerente ao tipo de processamento realizado, uma vez que a frequência de $[tʃiw]$ afetaria igualmente as duas sequências sonoras, demonstrando em ambos a preferência por acento paroxítono.

O favorecimento de $[V_{oral}.tʃiw]$ e $[V_{nasal}.tʃiw]$ ao acento paroxítono reforça a hipótese de que efeitos de frequência têm impacto nas representações fonológicas. Além disso, os resultados trazem à baila a hipótese de que a ausência de uma sequência sonora, $[niw]$ e $[V_{nasal}.tʃiw]$, com tonicidade

¹⁰ Ressalta-se que não é assumido que o processamento da palavra tenha ocorrido no nível silábico, visto que há indícios de que os informantes interpretavam as pseudopalavras como vocábulos, já que eles realizaram, em algumas palavras, processos fonológicos, como vozeamento de $[s]$ intervocálico, $[mu.za'niw]$ para $[mu.sa.niw]$, alçamento das vogais médias $[e]$ e $[o]$, $[ʃe.li.tor]$ e $[ka.pi'gor]$, e apagamento de $[r]$ em coda silábica, $[ba.zi'ga]$. Admite-se, contudo, que a tarefa de extração das sílabas, de certa forma, fez com que os falantes realizassem apenas abstrações silábicas para a composição das pseudopalavras, as quais, então, foram analisadas como palavras.

paroxítona diminuiria - mas não excluiria - a probabilidade de sua emergência (lacuna fonológica). Os dados de [niw] ilustram bem os efeitos da lacuna fonológica: o padrão acentual existente na língua, segundo o *Corpus* ABG, é predominante, [ˈniw] com 19 tipos e [ˈCiw] com 334 tipos, ao mesmo tempo em que padrões abstratos, como [ˈσ.Ciw] com 50 tipos, possibilitam o preenchimento das lacunas fonológicas por meio de processos de extensão analógica¹¹.

Por mais que a condição experimental possa ter direcionado a um processamento no nível silábico, as diferentes preferências acentuais, apontadas pelas análises estatísticas, de [niw] por acento oxítono e de [tʃiw] por acento paroxítono, fomentam questionamentos a respeito do papel do ataque silábico na atribuição de tonicidade. Note que a única informação que distingue esses dois tipos silábicos consiste no segmento que compõe o ataque silábico, os quais, como exposto anteriormente, possuem, juntamente com a rima, diferentes frequências. A problemática instaurada emerge da contraposição entre a proposta de Bisol (1994) e os resultados observados. Para a hipótese Métrica, o português seria sensível ao peso silábico, fato que colocaria a rima como constituinte central da sílaba, isto é, tanto o ataque quanto as informações segmentais não seriam consideradas pela proposta. Contudo, as diferentes proporções entre [niw] e [tʃiw] indicam que informações segmentais, tal como a do ataque silábico, e de frequência dos tipos silábicos também podem ser consideradas pelos falantes. Para a proposta Métrica, esses resultados estimulam o debate existente na literatura a respeito do papel que o ataque silábico desempenha na atribuição acentual (cf. Stuart, 1982). Para os Modelos Multirrepresentacionais, eles mostram a saliência do *template* silábico [CVG]¹² com tonicidade oxítona, mais especificamente do tipo [ˈtʃiw].

A *sílaba final* foi outra variável selecionada pelo primeiro modelo composto por [*grupo de análise* + *sílaba final* + *repetição*]. Observou-se que [ẽ.tʃiw] foi a única variante significativa, com $p = 0.0108$, a qual favorece o acento oxítono, como previsto por este estudo. A ausência de significância entre as variantes de *sílaba final* indica que os falantes não fazem distinção da vogal que precede

¹¹ Para mais detalhes, verifique a Figura 3.1 na pág. 76.

¹² G refere-se ao glide [w].

o tipo silábico final, exceto em [ẽ.tʃiw]. Esse fato reforça a justificativa de que os tipos silábicos finais, [niw] e [tʃiw], são realmente relevantes no momento da produção das pseudopalavras no experimento realizado, ao mesmo tempo em que incita questionamentos a respeito desse comportamento peculiar.

A significância de [ẽ.tʃiw] decorre da predileção das sequências [ẽ.tʃiw] e [ũ.tʃiw] por acento paroxítono em vocábulos como [po.lẽ.tʃiw], [ta.vẽ.tʃiw] e [xe.bũ.tʃiw], conforme demonstra a Tabela 3.7.

A amplitude de produções acentuais oxítonas que vai de 93 a 81 em [niw], de 79 a 65 em [V_{oral}.tʃiw] e de 77 a 58 em [V_{nasal}.tʃiw] ilustra a variabilidade de atribuições acentuais que pode ocorrer nas pseudopalavras, fato que era esperado, em vista da ausência de uso. Além disso, as diferentes produções acentuais nas pseudopalavras resultaram na seleção da variável *pseudopalavra* no segundo modelo estatístico [*pseudopalavra* + *repetição*]. A variabilidade dos resultados evidencia, ainda, que, embora a estrutura de palavra possua sílaba final pesada, as produções acentuais não são categóricas, como preveem as propostas algorítmicas. Ao contrário, há uma clara preferência pelo acento oxítono em decorrência da sua produtividade, porém, há também produções paroxítonas que demonstram preferências nítidas pelos tipos silábicos [V_{oral}.tʃiw] e [V_{nasal}.tʃiw].

A *repetição* também mostrou-se significativa nos dois modelos testados, [*grupo de análise* + *sílaba final* + *repetição*] e [*pseudopalavra* + *repetição*]. A sua seleção remete à condição experimental, posto que, como demonstra a Tabela 3.8, a quantidade de produções oxítonas aumentou conforme os falantes as repetiam, o que pode indicar que as pseudopalavras podem ter se comportado como gatilho acentual umas das outras, já que um único padrão acentual passou a ser produzido pelo falante.

Além da seleção de todas as variáveis discutidas acima, a variável que se mostrou mais significativa para a estabilidade dos modelos consistiu na variável aleatória *informante*, uma vez que as produções dos indivíduos foram simétricas em direção a um padrão acentual: oxítono, paroxítono ou paroxítono/oxítono. Em decorrência da complexidade e das discussões teóricas que os *informantes* incitam, reservou-se a seção 3.5 para a sua discussão.

De modo geral, investigou-se, no grupo de análise 1, em que medida

Tabela 3.7 Resultado de atribuições acentuais por *pseudopalavra* em [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw] (em ocorrências).

Pseudopalavra	Paroxítona	Oxítona
[bu.da.niw]	24	83
[mu.sa.niw]	16	90
[si.la.niw]	25	82
[ta.ʒe.niw]	21	85
[li.fe.niw]	20	87
[lo.pe.niw]	15	93
[ko.xu.niw]	16	91
[pe.tu.niw]	20	85
[tʃi.ku.niw]	25	81
[ko.ba.tʃiw]	38	65
[fe.ka.tʃiw]	33	74
[si.ma.tʃiw]	37	71
[go.le.tʃiw]	38	70
[su.pe.tʃiw]	32	75
[ve.ne.tʃiw]	35	73
[de.xu.tʃiw]	32	76
[na.tu.tʃiw]	27	79
[tʃi.ʒu.tʃiw]	37	68
[mo.sẽ.tʃiw]	29	77
[si.pẽ.tʃiw]	33	75
[ve.nẽ.tʃiw]	33	70
[ka.bẽ.tʃiw]	32	76
[ta.vẽ.tʃiw]	44	63
[po.lẽ.tʃiw]	49	58
[ba.pũ.tʃiw]	36	72
[bi.kũ.tʃiw]	35	68
[xe.bũ.tʃiw]	41	66

a frequência dos tipos silábicos pode influir na localização acentual. Por meio da concepção de lacuna fonológica, pressupôs-se que as pseudopalavras terminadas com os tipos silábicos [niw] e [V_{nasal}.tʃiw] receberiam principalmente tonicidade oxítona, enquanto o tipo [V_{oral}.tʃiw] seria produzido com acento paroxítono e oxítono. A hipótese de que o uso constrói e impacta na gramática fonológica do falante prediria que tipos silábicos não encontrados

Tabela 3.8 Resultado de atribuição acentual por *repetição* em [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw].

Repetição	Paroxítone	Oxítone
1 ^a	303	647
2 ^a	270	699
3 ^a	250	707

na língua tenderiam a não ser produzidos, em contrapartida, tipos silábicos recorrentes nas representações fonológicas teriam maior probabilidade de serem produzidos.

Os resultados demonstraram que os falantes tendem a atribuir mais frequentemente tonicidade paroxítone e oxítone aos padrões fonológicos que possuem os dois tipos acentuais. A semelhança no comportamento de [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], entretanto, indica que o desenho experimental desta pesquisa pode ter direcionado o processamento das unidades fonológicas ao nível silábico.

O grupo de análise 1, embora não apresente resultados simétricos aos previstos, traz evidências de que a frequência dos tipos silábicos constitui uma das inúmeras informações que, como num sistema complexo, combinadas podem prever a localização acentual no PB. Este resultado amplia resultados já encontrados para aspectos segmentais (cf. HUBACK, 2006; 2013; GARCIA & ZIMMER, 2010; HAUPT, 2011).

Ademais, este grupo de análise demonstrou que novos experimentos devem ser realizados, os quais precisam considerar fatores metodológicos e linguísticos não analisados neste experimento. Isso porque as abstrações, para além da sílaba, podem ser realizadas por meio do conteúdo semântico ou morfológico, identificando a sílaba como um todo e não apenas a rima [iw], as quais envolveriam padrões como *-átil*. A predição poderia ser verificada através de experimentos em que fossem contrastados pares de sentenças como *"Ele é muito* [ko.ba.tʃiw] *", que tenderia ao acento paroxítono, em oposição a "*Eu adoro [fe.ka.dʒiw] *", com propensão a receber acento oxítono. Ademais, novos testes podem ser realizados a fim de verificar se a vogal da sílaba precedente realmente não tem impacto na atribuição do acento em português. Um*

experimento em que pares de sentenças “*Ele é muito XX*[a.tʃiw]” e “*Ele é muito XX*[ẽ.tʃiw]” fossem contrastadas permitiria avaliar a relevância do tipo silábico da sílaba anterior. De qualquer maneira, ambas as propostas de experimento indicariam que a atribuição do acento tem relação com a morfologia e com o significado.

As ponderações acima permitem que estudos futuros sejam empreendidos para que se possa ter um conhecimento mais amplo sobre a relação entre tipos silábicos e atribuição do acento primário no PB.

3.3 Grupo de Análise 2: [-or] e [-ar]

Esta seção apresenta e discute os resultados obtidos no grupo de análise 2, o qual compreende as variantes [gor], [tor], [vor], [gar], [tar] e [var]. Este grupo investigou em que medida a produtividade dos tipos silábicos interagiria com os processos analógicos.

A hipótese, que surgiu a partir da emergência do acento paroxítono em [vor] no pré-teste¹³ e da falta de tipos silábicos com acento paroxítono no léxico da língua, buscava verificar se a alta frequência de ocorrência de palavras com similaridade fonológica às pseudopalavras engatilharia processos de extensão analógica entre a pseudopalavra-alvo e um vocábulo fonológico similar.

Esperava-se, com isso, que os tipos silábicos terminados em [-or], por possuírem semelhança fonológica com vocábulos paroxítonos frequentes¹⁴, receberiam um maior número de atribuições paroxítonas do que [-ar], o qual, além de ser altamente produtivo com tonicidade oxítona, tem poucos e infrequentes vocábulos com acento paroxítono¹⁵.

As hipóteses estabelecidas para o grupo de análise 2 são retomadas no Quadro 3.3:

¹³ Para mais detalhes, veja o anexo C.

¹⁴ Os tipos silábicos com acento paroxítono compreendem nomes próprios como *Igor*, *Vítor* e *Júnior*.

¹⁵ Os tipos silábicos com acento paroxítono compreendem vocábulos como *ímpar*, *Cortázar* e *Oscar*.

Grupo de Análise 2 [-or] e [-ar]	
[-or]	Tanto o acento oxítono quanto o paroxítono ocorrerão, uma vez que há palavras no PB terminadas em [-or] que são tanto oxítonas (ex. <i>pavor, fervor</i>) quanto paroxítonas (ex. <i>Vítor, Igor</i>), com maior incidência do acento paroxítono em decorrência da frequência dos vocábulos similares.
[-ar]	Tanto o acento oxítono quanto o paroxítono ocorrerão, posto que há palavras no PB terminadas em [-ar] tanto oxítonas (ex. <i>militar, cantar</i>) quanto paroxítonas (ex. <i>açúcar, Óscar</i>). Entretanto, o acento oxítono será preferido em decorrência da alta frequência de tipo de [-ar].

Quadro 3.3 Hipóteses do Grupo de Análise 2.

Como apresentado no Quadro 3.3, as hipóteses elaboradas para o grupo de análise 2 pressupõem que os tipos silábicos em [-or] tenderiam a receber mais atribuições acentuais paroxítonas em decorrência de processos extensionais entre as pseudopalavras e os vocábulos similares. Enquanto os tipos em [-ar] tenderiam a receber mais atribuições oxítonas em razão da alta produtividade com esse padrão acentual e da baixa frequência de ocorrência dos vocábulos similares com acento paroxítono. Essas predições podem ser atestadas ao contrastar a frequência de tipo dos tipos silábicos [-or] e [-ar] encontrados no *Corpus* ABG com os resultados obtidos no grupo de análise 2, como exibido na Tabela 3.9.

Tabela 3.9 Frequência de tipo das sequências sonoras [-or] e [-ar] por *tonicidade* no *Corpus* ABG e os resultados obtidos no Grupo de Análise 2.

Sequência Sonora	<i>Corpus</i> ABG		Resultados Obtidos ¹⁶	
	Paroxítona	Oxítona	Paroxítona	Oxítona
[-or]	9 (1,2%)	714 (98,8%)	204 (21,4%)	714 (74,8%)
[-ar]	14 (0,6%)	2.368 (99,4%)	148 (15,5%)	793 (82,9%)

$$\chi^2 = 21.695 (2) \text{ e } p < 0.0001$$

Considere a Tabela 3.9, a qual apresenta a frequência das sequências sonoras [-or] e [-ar] por *tonicidade* no *Corpus* ABG e os resultados obtidos no grupo de análise 2. Observe que, apesar de a frequência de [-or] e [-ar] com acento paroxítono ser similar, a frequência relativa dos tipos silábicos paroxítonos em relação ao mesmo tipo silábico com tonicidade oxítona é significativamente diferente. A diferença entre a quantidade dos tipos silábicos oxítonos faz com que [-ar], por ser um tipo de alta frequência, tenha conexões fonológicas mais fortes e arraigadas do que o tipo [-or], o qual, por ser de média frequência, é mais suscetível a processos analógicos (cf. BYBEE, 2010).

A predição é que a força das conexões fonológicas dos tipos [-or] e [-ar], decorrentes da frequência de tipo, motivaria diferentes comportamentos acentuais. Como exibido na Tabela 3.9, em ambos os tipos silábicos ocorreram produções tanto paroxítonas quanto oxítonas, como esperado. As diferentes proporções, no entanto, refletiram em preferências acentuais diferentes, i. é, paroxítona para [-or] e oxítona para [-ar], como será apresentado adiante. Isso ocorreria porque a média frequência do tipo [-or] com tonicidade oxítona faria com que suas conexões fossem menos resistentes e, com isso, mais suscetíveis a extensões analógicas do que [-ar], o qual por ter conexões fortes seria resistente a analogias.

Nos resultados gerais, conforme demonstra a Tabela 3.10, constatou-se a predominância do acento oxítono com 77,52% das ocorrências, seguido de 18,11% de paroxítonas e 2,62% de proparoxítonas. A diferença mostrou-se significativa com $\chi^2 = 4225.2$ (4) e $p < 0.0001$. A exclusão de 1,75% dos dados refere-se a produções silabificadas, com apagamento ou inserção de segmentos à pseudopalavra-alvo.

De maneira análoga ao grupo de análise 1, as hipóteses Métrica e Morfológica, embora prevejam a predominância do acento oxítono, se valeriam da extrametricidade e de uma regra de marcação a fim de justificar as produções proparoxítonas e paroxítonas. Como salientado anteriormente, tais propostas requerem uma marcação prévia que não poderia ser aplicada em vocábulos não reais, como é o caso das pseudopalavras. Quanto à hipótese Lexical,

¹⁶ As percentagens não contabilizam 100%, visto que foram obtidas 36 (3,8%) produções proparoxítonas no tipo [-or] e 15 (1,6%), no [-ar].

Tabela 3.10 Atribuição acentual no grupo de análise [-or] e [-ar].

Tonicidade	Ocorrências (%)
Proparoxítona	51 (2,62%)
Paroxítona	352 (18,11%)
Oxítona	1.507 (77,52%)
Excluídos	34 (1,75%)

$$\chi^2 = 4225.2 (4) \text{ e } p < 0.0001$$

além de propor uma atribuição aleatória - que não ocorre -, ao postular uma atribuição lexical, restringe a atribuição acentual a palavras reais. Em compensação, a hipótese Multirrepresentacional, até o momento, é a que melhor se adéqua aos resultados obtidos também no grupo de análise 2, posto que, como será exposto adiante, a frequência dos tipos silábicos figura como um fator relevante na atribuição acentual.

De acordo com os Modelos Multirrepresentacionais, a predileção pela tonicidade final reflete processos cognitivos de abstração e arraigamento dos padrões fonológicos, os quais fortificaram, por meio de instâncias de uso, os padrões produtivos, [-or] e [-ar]. Como pode ser visualizado na Figura 3.4, ao mesmo tempo em que as conexões abstratas foram se fortificando, ao projetarem suas conexões por similaridade fonológica, representadas na imagem pelas linhas cinzas, as conexões locais também foram tecendo relações associativas, sendo arraigadas às representações fonológicas, representadas pelas linhas marrons. A produtividade do padrão abstrato é visualizada na imagem por meio da espessura da nuvem de exemplar. Padrões mais produtivos possuem conexões mais grossas; padrões menos produtivos, conexões mais finas.

Figura 3.4 Esquema de conexões fonológicas entre [-or] e [-ar].

Para esse modelo, a competição entre os padrões produtivos e os padrões locais fomentaria a emergência do acento proparoxítono e do paroxítono, indicando que diferentes fatores, em diferentes níveis de abstração, contribuem para a determinação do acento em português. Não só a disputa entre padrões produtivos e locais atua no sistema, os vocábulos com similaridade fonológica também têm um papel relevante. Isso porque, conforme mostrou a Tabela 3.9, a atribuição acentual é, significativamente, diferente de acordo com o tipo silábico. As pseudopalavras do tipo [-or], ao serem produzidas, teriam suas conexões fonológicas estendidas aos esquemas que possuem vocábulos similares e maior índice de frequência. Em contrapartida, as pseudopalavras do tipo [-ar] teriam maior propensão a ser categorizadas com o padrão abstrato ['Car] em decorrência da sua alta produtividade e da baixa frequência dos vocábulos similares, resultando em menores índices de atribuições proparoxítonas e paroxítonas.

Tendo em vista que a frequência dos vocábulos fonológicos similares de [-or] e [-ar] consiste no fator que os distingue, os resultados do grupo de análise 2 indicam que a frequência dos tipos silábicos e os processos de extensão analógica atuam conjuntamente na determinação do acento em PB. O valor de significância, $p < 0.0001$, bem como a relação encontrada entre frequência

e padrão acentual nos tipos silábicos dos grupos de análise 1 e 2, minimiza a possibilidade de que esses resultados sejam aleatórios. As análises estatísticas, a serem apresentadas adiante, corroborarão tal afirmativa.

A emergência das tonicidades proparoxítonas e paroxítonas, ainda, contesta as predições algorítmicas que postulam sensibilidade ao peso silábico e as postulações de que haveria restrições na língua que impediriam a produção de proparoxítonos quando terminados em sílaba pesada. Para a hipótese de sensibilidade ao peso silábico, como afirmou Cantoni (2013), a constante correlação realizada entre estrutura de palavra e tonicidade reflete apenas uma herança do sistema acentual latino, o qual, por ser sensível ao peso silábico, fornece ao português uma gama de vocábulos que, de certa forma, ainda expressam a relação entre estrutura silábica e padrão acentual. Se o português fosse realmente sensível ao peso silábico, as produções de todas as pseudopalavras deste estudo seriam oxítonas, fato que não foi constatado. Para a impossibilidade de produções proparoxítonas em estruturas de palavra terminadas em sílaba pesada, a possível restrição do padrão acentual à estrutura de palavra será discutida em detalhes na seção 3.4.

O objetivo desta pesquisa ao criar o grupo de análise 2 foi demonstrar que, embora a produtividade dos tipos silábicos seja um fator importante na determinação do acento, quando se trata de pseudopalavras, as quais não têm frequência de ocorrência, processos analógicos com vocábulos similares de alta frequência também podem atuar. Para que essa predição fosse verificada, realizaram-se testes monofatoriais e multifatoriais entre a variável dependente, *tonicidade*, e as variáveis independentes, *grupo de análise*, *sílaba final*, *pseudopalavra* e *repetição*, bem como com a variável aleatória *informante*.

Diante da quantidade expressiva de produções proparoxítonas emergentes no grupo de análise 2 (51 ocorrências) em relação ao grupo 1 (2 ocorrências), decidiu-se incluí-las nas análises monofatoriais e multifatoriais com a utilização de um modelo de regressão logística multinomial. Contudo, ao correlacionar *tonicidade* com as demais variáveis, obteve-se uma grande quantidade de dados zerados, isto é, sem produções proparoxítonas, por isso, tanto as análises monofatoriais quanto as multifatoriais apresentaram desempenhos insatisfatórios. A falta de adequação dos modelos fez com que

somente as ocorrências paroxítonas e oxítonas fossem incluídas nas análises. A reorganização dos dados resultou em dados ainda significativos, como pode ser visualizado na Tabela 3.11.

Tabela 3.11 Resultados reestruturados de atribuição acentual em [-or] e [-ar] (sem proparoxítonas).

Tonicidade	Ocorrências (%)
Paroxítona	352 (18,83%)
Oxítona	1.507 (81,07%)
$\chi^2 = 38.614$ (1) e $p < 0.0001$	

As análises monofatoriais, por meio de testes de qui-quadrado (χ^2), inquiriram o papel que cada variável independente desempenhava na predição da variável dependente. Conforme demonstra a Tabela 3.12, somente as variáveis *grupo de análise* e *sílabas final* apresentaram valores significativos.

Tabela 3.12 Correlação entre a variável dependente e as variáveis independentes em [-or] e [-ar].

Variável	χ^2 (df)	Valor de p
<i>Grupo de Análise</i>	12.348 (1)	0.0004
<i>Sílabas Final</i>	14.315 (5)	0.01373
<i>Pseudopalavra</i>	20.496 (17)	0.2496
<i>Repetição</i>	0.56093 (2)	0.7554

De maneira análoga ao grupo de análise 1, as análises multifatoriais investigaram quais variáveis independentes interagiriam e prediriam o comportamento da variável dependente, *tonicidade*. A simetria entre os resultados dos dois grupos de análise fez com que o mesmo percurso estatístico empregado no grupo 1 fosse também realizado com o grupo 2.

Novamente, o modelo de regressão logística misto, tendo os informantes como variável aleatória, foi o modelo com melhor adequação e precisão. Os seguintes modelos foram testados: [*grupo de análise* + *sílabas final* + *repetição*] e [*pseudopalavra* + *repetição*], exatamente os mesmos modelos do grupo de análise 1. Os resultados completos dos modelos podem ser visualizados no anexo G.

O primeiro modelo, que testou [*grupo de análise + sílaba final + repetição*], selecionou todas as variáveis independentes, entretanto, apenas o *grupo de análise* mostrou-se significativo, com $p < 0.05$. Para as suas variantes, constatou-se que [-or] favorece o acento paroxítono, enquanto [-ar], o acento oxítono, conforme previam as hipóteses do grupo de análise 2. Esse resultado pode ser visualizado através das estimativas apresentadas na Tabela 3.13, nas quais os valores positivos, êxito, significam preferência oxítona, enquanto os valores negativos, fracasso, predileção paroxítona.

Tabela 3.13 Resultados do modelo de regressão logística com efeitos aleatórios em [-or] e [-ar] para grupo de análise.

	Estimativa	Erro Padrão	t	p
[-or]	-0.75627	0.36380	-2.079	0.0378*
[-ar]	3.34938	0.37173	9.010	< 2e-16***

Preferências, como de [-or] por acento paroxítono e de [-ar] por oxítono, não podem ser elucidadas pelas propostas algorítmicas. Isso porque para as hipóteses Métrica e Morfológica o acento nos vocábulos com esses tipos silábicos deveria se comportar de igual forma, i. é, com tonicidade oxítona. A postulação de regras categóricas ou de mecanismos de marcação desconsidera os efeitos de frequência e as informações segmentais que se mostram importantes para a determinação do acento no PB. Para os Modelos Multirrepresentacionais, as diferentes predileções refletem os efeitos de frequência dos padrões abstratos que estão arraigados às representações fonológicas. Dessa maneira, o favorecimento de [-or] pela tonicidade paroxítona espelha a frequência do tipo silábico [-or] e dos vocábulos que apresentam similaridades fonológicas, à medida que o favorecimento de [-ar] pelo acento oxítono retrata a produtividade do padrão [-ar] e a baixa frequência dos vocábulos similares. Ou seja, os Modelos Multirrepresentacionais permitem expressar tendências de comportamentos acentuais, as quais não são determinísticas.

Além disso, a diferença estatisticamente significativa entre as variantes [-or] e [-ar] no *grupo de análise* traz indícios de que efeitos de similaridade fonológica podem interagir com a frequência dos padrões fonológicos. A baixa

frequência de tipo de [-or] com tonicidade paroxítona, com 9 tipos, faz com que, segundo Bybee (2010), suas conexões fonológicas sejam fracas, tornando-o mais propício a sofrer processos analógicos com padrões de alta frequência. Diante de vocábulos frequentes, a analogia torna-se mais provável quanto mais similar forem as conexões fonológicas entre as pseudopalavras e seus vizinhos fonológicos. Nesse sentido, a semelhança entre *Vítor* e *Igor* com [fe.li.tor], [ka.pi.gor] e [zi.bi.tor] provoca a extensão do padrão acentual paroxítono a esses vocábulos. Em contrapartida, a preferência de [-ar] pelo acento oxítono decorre da sua alta produtividade em posição final, da baixa frequência desse tipo silábico com acento paroxítono e, ainda, da baixa frequência dos vocábulos fonológicos paroxítonos similares, fazendo com que o tipo silábico ['Car], por estar arraigado às representações, torne-se gatilho acentual do tipo [-ar].

É interessante notar que diferentes estudos na área da Psicologia investigaram a interação entre efeitos de similaridade lexical e de probabilidade fonotática. Shademan (2006), por exemplo, testou a interação entre analogia lexical (similaridade lexical) e conhecimento gramatical (probabilidade fonotática) em experimentos com diferentes estímulos: palavras reais e palavras reais/palavras novas. Os resultados revelaram efeitos interessantes. No experimento que se valia apenas de palavras reais, a probabilidade fonotática teve valores mais significativos do que a similaridade lexical. Em contrapartida, o experimento que contemplava palavras reais e palavras novas obteve resultado inverso: a similaridade lexical mostrou-se mais significativa do que a probabilidade fonotática. O estudo de Shademan (2006) fortalece os resultados encontrados no grupo de análise 2, posto que a preferência por produções paroxítonas em [-or] demonstra que a similaridade lexical atua mais fortemente em pseudopalavras que possuem vocábulos similares de alta frequência, embora seja possível também constatar efeitos de probabilidade fonotática através da maior quantidade de produções oxítonas em [-or] e [-ar].

A *sílabas final*, apesar de ser uma variável selecionada, não obteve valores significativos nos modelos de regressão logística em decorrência da simetria no comportamento dos tipos silábicos que compõem cada variante, como pode ser verificado na Figura 3.5. A preferência de [gar], [tar] e [var] por acento oxítono

em oposição a [gor], [tor] e [vor] por acento paroxítono reforça a hipótese de que a principal diferença entre os grupos consiste na sua terminação, ou *grupo de análise*.

Figura 3.5 Atribuição acentual por sílaba final em [-or] e [-ar].

O segundo modelo, que testou [*pseudopalavra* + *repetição*], selecionou todas as variáveis, mas apresentou valores significativos apenas para algumas pseudopalavras, como [no.vi.tar], [xa.pi.var] e [du.mi.tar], as quais favoreceram o acento oxítono. Em contrapartida, nenhuma pseudopalavra favoreceu o acento paroxítono, conforme a Tabela 3.14. Verifique os resultados completos do modelo de regressão logística com efeitos aleatórios no anexo G.

Tabela 3.14 Resultados parciais do modelo de regressão logística com efeitos aleatórios em [-or] e [-ar] para pseudopalavras.

	Estimativa ¹⁷	Erro Padrão	t	p
[no.vi.tar]	1.83786	0.65798	2.793	0.00527**
[xa.pi.var]	2.23249	0.67822	3.292	0.00102**
[du.mi.tar]	1.64437	0.63832	2.576	0.01007*

¹⁷ Lembre-se de que as estimativas com valores positivos, êxito, significam preferência oxítona, enquanto os valores negativos, fracasso, predileção paroxítona.

Embora o grau de informatividade das pseudopalavras, com 18 variantes, potencialize a probabilidade de a variável *pseudopalavra* ser significativa, a divergência na quantidade de atribuições acentuais paroxítonas entre as *pseudopalavras* do grupo [-ar] parece ser o fator motivador da significância desse grupo de vocábulos, como em [no.vi.tar], [xa.pi.var] e [du.mi.tar]. Em compensação, a menor discordância na quantidade de atribuições acentuais nas pseudopalavras terminadas em [-or] faz com que todas sejam igualmente relevantes, não apresentando quaisquer vocábulos significativos individualmente, como pode ser visualizado na Tabela 3.15. De maneira análoga ao grupo de análise 1, a amplitude de produções acentuais oxítonas em [-or] de 71 a 84 e de [-ar] de 78 a 91 ilustra a variabilidade de atribuições acentuais em pseudopalavras em decorrência desses vocábulos não possuírem frequência de uso.

Tabela 3.15 Atribuições acentuais por *pseudopalavra* em [-or] e [-ar] (em ocorrências).

Pseudopalavra	Paroxítona	Oxítona
[ka.pi.gor]	24	84
[me.dʒi.gor]	26	77
[po.si.gor]	23	84
[ka.mi.tor]	19	71
[fe.li.tor]	22	81
[zi.bi.tor]	24	81
[de.fi.vor]	21	86
[li.tʃi.vor]	25	80
[ta.ni.vor]	20	70
[ba.zi.gar]	22	78
[lu.fi.gar]	16	88
[se.tʃi.gar]	16	91
[ga.bi.tar]	21	84
[du.mi.tar]	16	91
[no.vi.tar]	14	91
[lo.ʒi.var]	14	89
[pe.si.var]	18	90
[xa.pi.var]	11	91

A *repetição* consistiu em outra variável selecionada nos dois modelos. Ao contrário do que ocorre no grupo de análise 1, ela não se mostrou significativa, posto que, como demonstra a Tabela 3.16, a quantidade de produções manteve-se relativamente constante entre as três repetições. O aumento nas atribuições oxítonas na terceira repetição decorre de uma mudança no padrão acentual, de proparoxítonas para oxítonas, em alguns vocábulos.

Tabela 3.16 Atribuições acentuais por *repetição* em [-or] e [-ar].

Repetição	Paroxítona	Oxítona
1 ^a	120	487
2 ^a	112	507
3 ^a	120	513

De modo semelhante ao grupo de análise 1, a variável aleatória *informante* foi inserida nos modelos estatísticos a fim de garantir estabilidade. A sua importância provém, novamente, dos diferentes comportamentos encontrados nas produções dos indivíduos, isto é, i. produções paroxítonas, ii. produções oxítonas e iii. produções paroxítonas e oxítonas. Em virtude de sua complexidade e dos debates que essa variável incita, ela será discutida em mais detalhes na seção 3.5.

O grupo de análise 2 investigou se a frequência dos tipos silábicos interagiria com efeitos de similaridade lexical em decorrência da frequência dos vocábulos fonológicos similares. Esperava-se, com isso, que a sequência sonora [-or], por ser um tipo infrequente com tonicidade paroxítona e apresentar vocábulos fonológicos similares frequentes com esse padrão acentual, receberia um maior número de atribuições acentuais paroxítonas do que a sequência [-ar], a qual, além de ser altamente produtiva com acento oxítono, possui poucos e infrequentes vocábulos fonológicos paroxítonos. Pressupôs-se, portanto, que a diferença na frequência dos vocábulos fonológicos de [-or] e [-ar] provocaria diferentes comportamentos entre essas sequências sonoras.

Os resultados demonstraram que os falantes atribuem uma maior quantidade de tonicidade paroxítona a vocábulos que possuem padrões fonológicos similares com alta frequência paroxítona, como ocorreu com [-or], do que os

padrões fonológicos com baixa frequência, como com [-ar]. Esses dados vão ao encontro do estudo de Shademan (2006), para a qual a similaridade lexical tem efeitos mais fortes do que a probabilidade fonotática em experimentos que envolvem palavras reais e palavras novas, posto que os vocábulos reais tornam-se gatilho analógico dos vocábulos não existentes.

Verificou-se, assim, que, embora a frequência dos padrões fonológicos impacte nas representações fonológicas, outros fatores podem interagir a depender da frequência dos tipos silábicos e dos vocábulos similares. Tipos silábicos infrequentes por terem conexões fonológicas fracas são propícios a sofrerem processos analógicos quanto maior a similaridade e a frequência dos vocábulos fonológicos similares. Em contrapartida, tipos silábicos frequentes têm conexões fonológicas fortes e resistentes a processos analógicos, já que a frequência de uso aumenta o nível de ativação e a rapidez no acesso lexical (cf. BYBEE, 2010).

Os resultados do grupo de análise 2 fomentam as evidências de que a frequência dos tipos silábicos impacta nas representações fonológicas e na maneira como o acento é produzido em pseudopalavras do PB. Além disso, mostram que os efeitos de frequência não ocorrem de forma independente de processos cognitivos, tal como a extensão analógica, os quais têm papel fundamental no preenchimento das lacunas distribucionais presentes nas redes e nos esquemas fonológicos dos falantes.

3.4 Proparoxítonas

Esta seção tem a finalidade de realizar uma análise qualitativa das proparoxítonas, contrapondo os dados empíricos obtidos neste experimento ao tratamento conferido a elas na literatura acentual. No PB, o acento proparoxítono tornou-se sinônimo de excepcionalidade, de casos marginais, de estrangeirismo, ou seja, ele é visto e tratado pelas análises linguísticas como um não padrão da língua. Autores como Bisol (1994) e Massini-Cagliari (1999), no âmbito da Fonologia Métrica, pressupõem que a sílaba final dos vocábulos proparoxítonos seria invisível às regras acentuais, as quais, segundo essa proposta, seriam sensíveis ao peso silábico. Lee (1995), por outro lado,

baseado na Fonologia Lexical, propõe que uma regra específica para palavras marcadas alteraria o algoritmo acentual de iambo para troqueu¹⁸.

Outros autores como Couto (1997) pressupõem, ainda, que regras fonotáticas impediriam a ocorrência de acento proparoxítono em vocábulos que possuem a última e/ou a penúltima sílabas pesadas. Exceções são encontradas e apontadas pela autora, contudo, elas pertenceriam ao "nível erudito de linguagem" ou constituiriam "casos periféricos" da língua. As proparoxítonas seriam assim, pois, segundo Couto (1997), os falantes tenderiam a desfazê-las, como ['awku] para *álcool* e ['fasi] para *fácil*, já que haveria uma vogal subjacente em vocábulos terminados em /r/ e /l/, tal como /'util/ para *útil*. A postulação da autora assemelha-se à consoante subjacente prevista por Bisol (1994), a qual só se manifestaria em formas derivadas, como em *alcoholismo* e *alcoholizar*¹⁹.

Opondo-se à marginalização das proparoxítonas, o estudo de Araújo et al (2007) refuta os principais argumentos empregados por esses e outros autores a fim de colocar as proparoxítonas à margem das análises acentuais. Para os autores que defendem uma entrada tardia desses vocábulos no léxico do português, já que processos de redução fonológica, anteriores ao século XVI, teriam alterado a localização acentual de proparoxítono para paroxítono, Araújo et al (2007) mostram que as proparoxítonas, bem como as não proparoxítonas, tiveram uma entrada regular na língua, com picos nos séculos XII, XVI e XIX para todos os padrões acentuais. Já para aqueles que defendem que as proparoxítonas estão sujeitas a processos fonológicos, como a síncope, que reduz os vocábulos com a finalidade de desfazer o padrão proparoxítono, os autores sustentam que, aproximadamente, 62,7% formariam *clusters* que violariam as restrições fonotáticas da língua, como em ['mɛ.dʒi.ku] > *['mɛ.dku] e ['be.ba.du] > *['be.bdu]²⁰. Ainda, para as afirmativas de que os falantes evitam o uso das proparoxítonas, eles demonstram, por meio de uma análise de frequência, que a sua baixa produtividade decorre do tamanho médio desses vocábulos, posto que há uma relação direta entre frequência de uso e

¹⁸ Veja mais detalhes dessas propostas na seção 1.3.

¹⁹ Exemplos extraídos de Couto (1997, p. 136).

²⁰ Dados extraídos de Araújo et al (2007, p. 48-49).

extensão das palavras, isto é, quanto maior a palavra menor a probabilidade de sua ocorrência.

A defesa de Araújo et al (2007) para a manutenção das proparoxítonas nos estudos acentuais é legitimada por estudos experimentais como o desta dissertação. A ocorrência de 53 (1,1%) produções proparoxítonas, produzidas por 16 falantes (quase 50% dos indivíduos), configura-se como uma evidência de que o padrão proparoxítono pertence ao sistema acentual da Língua Portuguesa. A baixa produtividade das proparoxítonas não deve ser compreendida como a importação de um padrão estrangeiro, entretanto, como a manifestação de um padrão que está presente no conhecimento gramatical dos falantes e que tende a se manifestar mais frequentemente em certos contextos do que em outros.

Estudos têm que ser realizados a fim de mapear o comportamento das proparoxítonas dada a complexidade do sistema acentual do português, todavia, é possível verificar em *corpora* linguísticos, por exemplo, as previsões de Couto (1997) de que as proparoxítonas não podem ter qualquer uma das duas sílabas finais pesadas. No *Corpus* ABG, verifica-se a estrutura de palavra CV-CV-CVC entre as 20 mais frequentes com acento proparoxítono, com 42 ocorrências. Além disso, os resultados deste experimento contribuem para a rejeição dessa "restrição fonotática" do PB, já que foram obtidas 53 produções proparoxítonas também com a estrutura de palavra CV-CV-CVC.

Os dados deste estudo também trazem evidências de que, embora a relação entre estrutura de palavra e padrão acentual pareça ser direta, há sim outros fatores que interagem e contribuem para a determinação do acento em português. Por mais que haja uma propensão a tecer relações categóricas entre estrutura de palavra X e padrão acentual Y, essas relações devem ser interpretadas como tendências de um sistema linguístico formado por meio de abstrações emergentes de experiências de uso. É importante lembrar que uma grande parcela do léxico do português provém da língua latina, a qual era sensível ao peso silábico, por isso, é natural que as redes de conexões fonológicas expressem a relação entre estrutura de palavra e padrão acentual (cf. CANTONI, 2013).

A estrutura consiste em um dos fatores relevantes para o acento, contudo,

como este estudo demonstrou, os tipos silábicos emergem como um fator robusto para a predição do acento, através de conexões que se fortalecem em diferentes níveis, do segmental ao estrutural. A influência de traços segmentais e da frequência de uso foi visualizada já nos resultados do metaexperimento (veja a seção 2.1.3), no qual o tipo silábico [sĩ], por apresentar alta frequência com tonicidade proparoxítona, em 16 de 27 tipos tônicos, emergiu em 6 das 9 produções do vocábulo [sĩ.do.to]. Neste experimento, verifica-se que as proparoxítonas manifestam-se, majoritariamente, em [-or], conforme a Tabela 3.17, mais especificamente nas pseudopalavras [ta.ni.vor] e [ka.mi.tor], veja a Tabela 3.18.

Tabela 3.17 Atribuições acentuais proparoxítonas por *grupo de análise*.

Grupo de Análise	Proparoxítona	Paroxítona	Oxítona
[-ar]	15	148	793
[-or]	36	204	714
[niw]	1	182	777
[V _{oral} .tʃiw]	1	309	651
[V _{nasal} .tʃiw]	0	332	625

Embora a quantidade de produções proparoxítonas em [ta.ni.vor] e [ka.mi.tor] seja relativamente alta e o contraste entre o grupo de análise 1 e 2 coloque questões importantes para os estudos acentuais, não é possível tecer uma análise de frequência de tipos, já que não houve um controle prévio da frequência dessas sequências sonoras. Estudos futuros direcionados à investigação do comportamento das proparoxítonas podem elucidar esses resultados.

O presente estudo, portanto, traz como uma de suas contribuições a constatação de que as proparoxítonas pertencem ao sistema acentual do português, estando presentes nas representações mentais dos falantes, de forma que a sua manifestação em pseudopalavras torna-se apenas a extensão de um padrão fonológico possível da língua. Além disso, há evidências de que o acento proparoxítono pode incidir em vocábulos que possuem sílaba final pesada, embora com baixo índice de ocorrência.

Tabela 3.18 Atribuições acentuais proparoxítonas por *pseudopalavra* (em ocorrências).

Grupo de Análise	Pseudopalavra	Proparoxítona	Paroxítona	Oxítona
[-ar]	[ba.zi.gar]	3	22	78
	[du.mi.tar]	1	16	91
	[lo.ʒi.var]	3	14	89
	[lu.fi.gar]	1	16	88
	[xa.pi.var]	6	11	91
	[se.tʃi.gar]	1	16	91
[-or]	[ka.mi.tor]	16	19	71
	[zi.bi.tor]	3	24	81
	[fe.li.tor]	4	22	81
	[me.dʒi.gor]	2	26	77
	[ta.ni.vor]	11	20	70
[-iw]	[tʃi.ku.niw]	1	25	81
	[na.tu.tʃiw]	1	27	79

3.5 Informante

Esta seção destina-se à discussão do papel desempenhado pela variável aleatória *informante*, a qual foi inserida nas análises a fim de garantir a estabilidade e a adequação dos modelos estatísticos. Observe a Tabela 3.19, na qual é possível visualizar e compreender como os *informantes* atribuíram tonicidade às pseudopalavras. Note que, como apontaram as análises estatísticas, os indivíduos apresentam comportamentos hierarquizados por *clusters* em três diferentes direções: i. informantes que tendiam a atribuir acento oxítono; ii. informantes que tendiam a atribuir acento paroxítono; e iii. informantes que tendiam a atribuir "igualmente" acento paroxítono e oxítono.

Tabela 3.19 Atribuições acentuais por *informante*.

Grupo	Informante	Proparoxítona	Paroxítona	Oxítona
i	A.O.	1	14	119
	A.C.	0	2	130
	C.C.	0	2	132
	C.U.	0	0	135
	A.L.	0	16	119
	D.G.	1	15	116
	D.N.	4	16	112
	E.G.	1	1	114
	F.S.	0	4	147
	G.N.	3	3	128
	G.R.	0	13	121
	H.P.	0	5	125
	I.P.	0	1	134
	I.B.	0	2	130
	L.M.	0	4	130
	M.J.	0	4	127
	P.R.	0	0	135
	R.A.	1	1	129
	R.S.	0	2	132
	R.F.	0	3	131
	R.S.	2	2	125
	V.A.	0	8	127
ii	G.R.	5	110	16
	G.B.	0	135	0
	J.T.	7	122	6
	L.A.	0	117	16
	T.A.	2	112	10
	T.S.	1	121	12
iii	A.T.	7	30	98
	C.D.	0	31	104
	J.C.	11	73	49
	K.C.	1	74	60
	L.R.	0	20	115
	M.C.	3	26	99
	N.G.	0	21	110
	T.A.	3	65	67

A sistematicidade de atribuições acentuais dos indivíduos nos grupos i e ii sugere que os falantes interpretam semelhantemente os tipos silábicos, assumindo um mesmo padrão acentual para todas as pseudopalavras. Em contrapartida, no grupo iii, a consciência da frequência dos padrões fonológicos parece realmente existir, posto que as atribuições acentuais seguem as hipóteses discutidas ao longo deste trabalho, como mostra a Tabela 3.20.

Tabela 3.20 Atribuições acentuais do grupo iii por *grupo de análise*.

Grupo de Análise	Proparoxítona	Paroxítona	Oxítona
[-ar]	4	14	197
[-or]	20	53	140
[niw]	1	36	175
[V _{oral} .tʃiw]	0	114	102
[V _{nasal} .tʃiw]	0	123	88

A relevância e a hierarquização do comportamento dos *informantes* suscitam difíceis e importantes questionamentos a respeito do papel do indivíduo nos estudos acentuais, visto que, embora existam estudos que abordem a variação no nível segmental, individual ou social, quer na Fonologia (cf. CRISTÓFARO-SILVA & OLIVEIRA, 2001; BATTISTI, 2007; ABAURRE, 2013), quer na Sociolinguística (cf. CALLOU, MORAES & LEITE, 1998; MENDES & OUSHIRO, 2012; OUSHIRO, 2015), encontraram-se somente estudos para a língua russa que abordam a variação intra e inter-falantes no nível acentual (cf. PIPERSKI & KUKHTO, 2016; LAGERBERG, 2007).

Diferentes fatores são apontados na literatura russa como possíveis motivadores das variações acentuais: variação sociolinguística, processos analógicos, frequência das palavras, influência de jargões profissionais, semântica e origem etimológica do vocábulo. Dentre eles, apenas os dois primeiros agentes poderiam ser aplicados às pseudopalavras.

Para a variação social, características sociais como idade, sexo, escolaridade, nível socioeconômico e profissão consistem nos principais fatores indicados nos estudos Sociolinguísticos (cf. LABOV, 2008). Neste estudo, entretanto, esses aspectos não parecem influir na atribuição acentual. Como exposto na Tabela 3.21, de maneira geral, todos os grupos possuem indivíduos

com semelhantes características sociais, isto é, idade, gênero, naturalidade e, principalmente, escolaridade.

Tabela 3.21 Informações sociais de uma amostra de informantes.

Grupo	Informante	Idade	Gênero	Naturalidade	Curso Superior
i	C.U.	27	homem	Santos (SP)	Engenharia
	I.P.	19	mulher	São Paulo	Letras
	H.P.	29	mulher	Fortaleza (CE)	Engenharia
	M.J	58	mulher	Orlândia (SP)	Psicologia
ii	G.B.	19	mulher	Bilac (SP)	Letras
	G.R.	20	homem	São Paulo	Letras
	T.S.	49	mulher	São Paulo	Gestão de Pessoas
iii	C. D.	25	homem	São Paulo	História
	J.C.	22	homem	Capivari (SP)	Letras
	M.C.	61	mulher	São Paulo	Contabilidade

Quanto aos processos analógicos, eles, realmente, justificam o grupo iii, em que, como salientado ao longo desta pesquisa, a frequência dos tipos silábicos e dos vocábulos similares tende a influir na maneira como o acento é produzido. Nos grupos i e ii, todavia, a produção sistemática dos indivíduos inviabiliza semelhante justificativa. Eles, na verdade, parecem assumir um padrão acentual e aplicá-lo a todos os vocábulos, seja o padrão paroxítono, seja o oxítono, mas, quando variam, o padrão variante tende a seguir as predições deste estudo, como ocorreu com o falante G.R., em que, das 13 produções paroxítonas, todas foram em [tʃiw].

O fato de ser um experimento e das pseudopalavras não terem frequência de uso tem que ser considerado, o que pode ter influenciado no comportamento dos falantes. Apesar disso, os resultados deste estudo contribuem para evidenciar que, embora com produções simétricas nos grupos i e ii, as variações acentuais apresentam as tendências visualizadas no grupo iii e, de maneira mais ampla, os resultados totais desta pesquisa. Além disso, ressaltam a importância de se fazer estudos no âmbito acentual que passem a considerar o informante como uma variável relevante e que façam desenhos experimentais que investiguem o papel dos indivíduos. A explicativa dos diferentes comportamentos acentuais,

portanto, permanece aberta para investigações futuras.

3.6 Considerações Finais

O presente capítulo apresentou e discutiu os resultados obtidos em um desenho experimental que investigou se a frequência dos tipos silábicos influi na determinação do acento primário em PB. Para que essa hipótese fosse verificada com mais precisão, definiram-se dois grupos de análise.

O primeiro grupo, composto pelas variantes [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], inquiriu, por meio do conceito de lacuna fonológica, em que medida a frequência dos tipos silábicos pode prever a localização acentual. Os resultados demonstraram que os falantes fazem abstrações de padrões fonológicos para atribuir acento às pseudopalavras. Além disso, a falta de distinção da vogal precedente de [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], possivelmente influenciada pela tarefa experimental, demonstra que a sílaba consiste no elemento considerado pelos falantes para a atribuição acentual.

O segundo grupo, formado pelas variantes [-or] e [-ar], analisou se a frequência dos tipos silábicos interagiria com efeitos de similaridade lexical em decorrência da frequência dos vocábulos fonológicos similares. Os resultados apontaram que tanto a similaridade lexical quanto a produtividade dos padrões fonológicos são relevantes. Em sequências sonoras com alta produtividade e com vocábulos similares com baixa frequência, os efeitos de frequência (produtividade) são mais fortes do que os de similaridade (analogia), em contrapartida, quando os vocábulos similares são frequentes e o tipo silábico é de baixa ou média frequência, os efeitos de similaridade lexical se sobrepõem aos efeitos de frequência.

Os resultados deste estudo indicam que diferentes fatores, em diferentes níveis de complexidade, contribuem para a determinação da localização acentual, tais como a frequência dos tipos silábicos e os processos analógicos, demonstrando que o sistema acentual do português apresenta tendências que podem ser expressas por meio de efeitos de frequência. Além disso, a emergência do acento proparoxítono traz evidências que asseguram este padrão acentual como pertencente ao sistema fonológico do PB, posto que

proparoxítonos são produzidos até mesmo em pseudopalavras. Por fim, os resultados demonstram ainda que mais investigações devem ser realizadas a fim de compreender o papel e os diferentes comportamentos dos *informantes*.

4

Conclusão

Esta dissertação teve o objetivo de investigar o papel que a frequência dos tipos silábicos desempenha na atribuição acentual em pseudopalavras no português brasileiro. A partir dos Modelos Multirrepresentacionais, contrapuseram-se teorias fonológicas a dados empíricos a fim de avaliar a complexidade do sistema acentual do português.

Para isso, apresentaram-se as principais propostas de análise para o acento primário no PB. Pautadas em perspectivas fonológicas ou morfológicas, as propostas fazem uso de algoritmos para justificar as simetrias encontradas na língua, aos casos assimétricos, entretanto, restam pretextos de excepcionalidade. De modo geral, as análises buscam um único fenômeno regulador e preditor do sistema acentual. Em um sistema tão complexo como a língua, supor que um único agente regule todo um sistema acentual é simplificar a complexidade da língua e atribuir a ela uma natureza categórica, determinística.

Contrapondo-se às perspectivas tradicionais, assumiu-se, ancorada nos Modelos Multirrepresentacionais, uma concepção de língua não reducionista em que as experiências de uso conjugam-se aos processos cognitivos gerais a fim de que a língua emergja. Nessa perspectiva, a gramática é formada por meio de uma série de processos de abstração, fortificação e arraigamento que partem dos usos concretos para a formação de esquemas complexos e abstratos. Ao conceber a língua como um sistema dinâmico, admitiu-se que diversos agentes, em diferentes níveis hierárquicos, são responsáveis pela atribuição acentual no PB.

Dentre as propostas acentuais analisadas, a hipótese Multirrepresentacional, de Cantoni (2013), por ser mais abrangente, prevê que diferentes padrões acentuais, como observado neste experimento, possam ocorrer. Para a autora, o estudo de um sistema linguístico deve partir da observação de estágios anteriores a ele, já que nos Sistemas Dinâmicos o tempo consiste num preditor importante para o atual estado de um sistema. Nesse sentido, a língua-mãe transmitiria às línguas descendentes algumas de suas características principais. Por o português ser proveniente do latim, ele teria herdado os seus *templates* acentuais - proparoxítono, paroxítono e oxítono -, mas não a sensibilidade ao peso silábico, que regulava o sistema acentual latino, já que ela teria sido perdida na passagem do latim para o português com a perda da quantidade vocálica.

Embora tenha sofrido diversos processos fonológicos e morfológicos ao longo do tempo, a herança de um léxico que ainda apresentava resquícios da sensibilidade ao peso silábico fez com que diversas análises linguísticas admitissem o português como sensível ao peso silábico. Contudo, como evidenciaram os dados deste estudo, essa aparente sensibilidade deve ser interpretada apenas como uma tendência do sistema fonológico, já que, como afirmou Cantoni (2013), a inserção dos vocábulos latinos na Língua Portuguesa provocou a emergência de estruturas abstratas que, de certa forma, expressam a correlação entre estrutura de palavra e padrão acentual.

Trata-se de uma tendência, visto que outras línguas também contribuíram para a formação do léxico do PB, especialmente as línguas indígenas, africanas, o grego e o inglês, as quais forneceram padrões fonológicos divergentes do padrão latino. Enquanto, de um lado, o latim introduziu a estrutura de palavra como uma variável relevante, do outro lado, as línguas indígenas, por exemplo, podem ser responsáveis pela importância das informações segmentais, já que alguns vocábulos terminados em *i* e *u* provêm do tupi e têm acento final. Entende-se, portanto, que a interação entre diferentes padrões fonológicos, provenientes das diversas línguas que compõem o léxico do português, foi responsável pela formação de redes complexas e abstratas, as quais são consolidadas através do uso.

Ao admitir que o uso impacta na gramática dos falantes, pretendeu-se

com este estudo inquirir como a frequência de tipo, a qual é responsável pela produtividade dos padrões fonológicos, pode prever a localização acentual. Para que a investigação pudesse ser feita, fez-se necessária a realização de duas etapas preliminares.

A primeira etapa consistiu na compilação de um *corpus* linguístico por meio do qual pudesse ser extraída a frequência dos vocábulos e dos padrões fonológicos. O *Corpus* ABG, como foi denominado, reuniu diferentes gêneros de textos, orais e escritos, sistematizados segundo os critérios determinados pela Linguística de *Corpus*. Além da seleção e compilação dos textos, utilizaram-se e criaram-se ferramentas computacionais, como silabificador, transcritor e acentuador fonológico, a fim de facilitar a manipulação e a extração dos dados. Todos os instrumentos foram desenvolvidos em Linguagem Python.

A segunda etapa consistiu na elaboração de uma metodologia que pudesse elicitar os padrões acentuais sem que os estímulos e o desenho experimental interferissem nos resultados. A verificação do modo de apresentação dos estímulos fez-se necessária, visto que a Língua Portuguesa utiliza diacríticos para marcar a sílaba tônica em alguns vocábulos. Por isso, um metaexperimento, como foi denominado, teve que ser desenvolvido, o qual comparou diferentes formas de apresentação dos estímulos - auditivos, imagéticos e ortográficos -, para que a condição experimental mais adequada fosse utilizada.

Os resultados do metaexperimento, o qual identificou a condição imagem como a mais adequada, tornaram-se os primeiros indícios de que a frequência dos tipos silábicos também é considerada pelos falantes no momento da atribuição acentual. A incidência de diferentes padrões acentuais em [niw] e [tʃiw], por exemplo, fortaleceu a hipótese de que a frequência impacta na maneira como o acento é produzido, fato que conduziu à adoção de [niw], [V_{oral}.tʃiw] e [V_{nasal}.tʃiw] para compor o grupo de análise 1.

A predição do grupo de análise 1 de que a ausência de tipos fonológicos com um determinado padrão acentual no léxico da língua diminuiria a probabilidade de sua emergência foi atestada nos tipos [niw] e [V_{oral}.tʃiw], mas não em [V_{nasal}.tʃiw]. A preferência de [niw] pelo acento oxítono e de [V_{oral}.tʃiw] pelo acento paroxítono, obtida nos testes de regressão logística, constitui uma evidência a favor da hipótese deste estudo. O tipo silábico [V_{nasal}.tʃiw], en-

tretanto, demonstrou preferência paroxítona em vez de oxítona, contrariando as hipóteses iniciais.

A justificativa atribuída ao comportamento inesperado de [V_{nasal}.tʃiw] pautou-se nos estudos de Vitevitch et al (1997), para o qual palavras não reais são processadas por vias sublexicais, em diferentes n-gramas (bigramas, trigramas). Ao tecerem diferentes abstrações às pseudopalavras, os falantes, provavelmente, limitaram-se ao nível silábico, posto que as instruções da tarefa solicitavam que os indivíduos extraíssem o primeiro som (i. e., da primeira sílaba) da palavra, o que, de certa forma, pode ter conduzido a um processamento silábico.

A constatação de que a frequência dos tipos silábicos consiste em um dos fatores que predizem a localização acentual, somada ao comportamento semelhante entre [V_{oral}.tʃiw] e [V_{nasal}.tʃiw], indica que o acento pode considerar a sílaba como unidade de análise. Além disso, os resultados trouxeram evidências de que as informações segmentais também são acessadas pelos falantes e não somente a estrutura de palavra.

A escolha do grupo de análise 2 partiu do pré-teste, o qual atestou produções paroxítonas em tipos silábicos que não possuíam esse padrão acentual no léxico da língua, como [vor]. As ocorrências paroxítonas em [vor] foram contrastadas com a frequência do tipo [-or] e dos vocábulos similares, atestando-se que, embora o tipo apresentasse média frequência, possíveis efeitos de similaridade lexical poderiam ter ocorrido, já que os vocábulos similares eram mais frequentes. Para contrapor o tipo silábico [-or], selecionou-se a sequência sonora [-ar], a qual apresenta alta produtividade com acento oxítono e vocábulos paroxítonos com baixa frequência de uso.

O grupo de análise 2, composto por [-or] e [-ar], inquiriu se a produtividade dos padrões fonológicos poderia interagir, por meio de processos analógicos, com os vocábulos similares dos tipos silábicos. Como as hipóteses previam, os resultados demonstraram que tipos silábicos com vocábulos similares frequentes, [-or], eram mais propícios a sofrer processos analógicos do que tipos silábicos com alta produtividade e vocábulos similares infrequentes, [-ar], já que a frequência fortalece as conexões locais, diminuindo a probabilidade de mudanças analógicas.

Dessa maneira, os resultados dos grupos de análise 1 e 2 demonstraram que a frequência dos tipos silábicos impacta na forma como o acento é produzido no PB, bem como trouxeram evidências de que processos analógicos podem interagir com a produtividade na atribuição acentual. Além disso, a emergência do acento proparoxítono nos dois grupos de análise tornou-se uma evidência do seu pertencimento ao sistema acentual do português e de que, aparentemente, não há restrições de estruturas de palavra que impeçam a sua ocorrência. Ao contrário, há padrões abstratos produtivos que favorecem este ou aquele padrão acentual, mas não regras categóricas.

A perspectiva de análise adotada, portanto, concebeu a língua como fruto de instâncias de uso e de processos cognitivos, que interagem e constroem padrões robustos que são expressos por meio de tendências linguísticas não determinísticas. A língua deixa de ser estável e categórica e assume uma face dinâmica. A dinamicidade da língua faz com que diferentes agentes sejam responsáveis pela emergência de um padrão fonológico, posto que suas redes fortes, redundantes e salientes se tornarão gatilho para os padrões novos da língua.

Ademais, assumiu-se que o acento nas palavras reais, como apontou Cantoni (2013), deve ser lexicalmente especificado, posto que é adquirido, simultaneamente, ao aprendizado das palavras. Com relação aos vocábulos não reais, ou pseudopalavras, restam aos esquemas fonológicos tecer relações associativas entre estrutura de palavra, produtividade dos padrões fonológicos e similaridade lexical com vocábulos reais a fim de construir as conexões entre vocábulos reais e novos vocábulos.

Em suma, esta pesquisa traz como uma de suas contribuições a demonstração de que a frequência dos tipos silábicos consiste em uma das inúmeras variáveis que contribuem para a predição do acento primário em PB, juntamente com os efeitos de similaridade lexical, ou processos analógicos. Além disso, verificou-se que os falantes utilizam os mesmos correlatos fonéticos empregados nas palavras reais para atribuir tonicidade às pseudopalavras. Este trabalho também constatou a necessidade de se realizar novos estudos a fim de mapear o comportamento do acento proparoxítono e a revelância dos indivíduos nos estudos acentuais.

Referências Bibliográficas

ABAURRE, M. B. M. (org.) *A construção fonológica da palavra: Gramática do português culto falado no Brasil*. São Paulo: Contexto, 2013. 252 p.

ALBANO, E. C. O Gesto e Suas Bordas: esboço de Fonologia Acústico-Articulatória do Português Brasileiro. *Mercado das Letras*, São Paulo, 2001. 272 p.

ARAÚJO, L. C. *Statistical Analyses in Language Usage*. 2013. 199 f. Tese (Doutorado em Engenharia Elétrica). Escola de Engenharia, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

ASPA. *Projeto ASPA: Avaliação Sonora do Português Atual*. Disponível em: <<http://www.projetoaspa.org/busca2>>. Acesso em 19/12/2014.

Audacity Team. *Audacity(R): Free Audio Editor and Recorder* [Programa de Computador]. Versão 2.1.2. Disponível em <<http://www.audacityteam.org/>>.

BARBOSA, J. M. *Introdução ao Estudo da Fonologia e Morfologia do Português*. Livraria Almedina, Coimbra, 1994. 295 p.

BATTISTI, E. et al. Palatalização das oclusivas alveolares e a rede social dos informantes. *Revista Virtual de Estudos da Linguagem. ReVEL*. Vol. 5, n. 9, agosto de 2007. p.1-29.

BECHARA, Evanildo. *Moderna Gramática Portuguesa*. Rio de Janeiro: Lucerna, 2005. 37. ed. rev. e ampl. 15. reimp.

BECKMAN, M. E. & KINGSTON, J. Introduction, papers in Laboratory Phonology I: between the grammar and physics of speech (reprint). In: COHN, A. C.; FOUGERON, C. & HUFFMAN, M. K. *The Handbook of Laboratory Phonology*. Oxford University Press. New York, 2012. p. 10- 16.

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.
- BENEVIDES, A. L. *Flutuação do Acento em Palavras Produzidas por Falantes Nativos do Português Brasileiro*. 2013. 91 f. Relatório de Pesquisa. Departamento de Linguística, Universidade de São Paulo, São Paulo, 2013.
- BENEVIDES, A. L. O acento oscilante em português brasileiro e europeu. *Cadernos do IL*, Porto Alegre, n.º 50, junho de 2015. p. 28-47.
- BERENT, I. Pronological Priming in the Lexical Decision Task: regularity effects are not necessary evidence for assembly. *Journal of Experimental Psychology: Human Perception and Performance*, v. 23, n. 6, p. 1727-1747, 1997.
- BIBER, D. Representativeness in Corpus Design. *Literary and Linguistic Computing*, v. 8, n. 4, 1993. p. 243-257.
- BISOL, L. O acento e o pé métrico. *Letras de Hoje*, Porto Alegre, v. 29, n. 4, p. 25-36, dezembro 1994.
- BISOL, L. A sílaba e seus constituintes. In: ABAURRE, M. B. M. (org.) *A construção fonológica da palavra*. São Paulo: Contexto, 2013. v. III. p. 21-52.
- BOB, R.; HAY, J. & JANNEDY, S. *Probabilistic Linguistics*. Cambridge: The MIT Press, 2003. 451 p.
- BOERSMA, P. & WEENIK, D. *Praat: doing phonetics by computer [Computer program]*. Version 5.3.51, retrieved 2 June 2013. Disponível em: <<http://www.praat.org/>>.
- BYBEE, J. *Morphology: a study of the relation between meaning and form*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1985. 246 p.
- BYBEE, J. *Phonology and Language Use*. Cambridge: Cambridge University Press, 2001. 238 p.
- BYBEE, J. From usage to grammar: the mind's response to repetition. Project MUSE: Scholarly journals online. 2006. p. 711-732.
- BYBEE, J. *Language, Usage and Cognition*. Cambridge: Cambridge University Press, 2010. 252 p.
- BYBEE, J. & BECKNER, C. Emergence at the crosslinguistic level:

attractor dynamic in language change. In: MACWHINNEY, B. & O'GRADY, W. (eds). *The Handbook of Language Emergence*. Wiley Backwell, 2015. p. 183-200.

BYBEE, J. & HOPPER, P. (eds.). *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001. 494 p.

CAGLIARI, L. C. *An experimental study of nasality with particular reference to Brazilian Portuguese*. Tese (Doutorado em Linguística). University of Edinburgh, Edinburgo. 1977.

CALLOU, D. MORAES, J. & LEITE, Y. Apagamento do R final no dialeto carioca: um estudo em tempo aparente e em tempo real. *DELTA*, v. 14, 1998. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-44501998000300006>.

CÂMARA Jr., J. M. *Estrutura da Língua Portuguesa*. Petrópolis: Editora Vozes. 34. ed. 2001 [1970].

CANTONI, M. M. *O Acento no Português Brasileiro: uma abordagem experimental*. Tese (Doutorado em Linguística Teórica e Descritiva). Estudos Linguísticos da Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2013.

CAPOVILLA, F. C. & CAPOVILLA, A. G. S. Problemas de aquisição de leitura e escrita: efeitos de déficit de discriminação fonológica, velocidade de processamento e memória fonológica. *Estudos e Pesquisas em Psicologia*, UERJ, RJ, ano 2, n. 1, 1º semestre de 2002. p. 25-50.

CAPOVILLA, F. C.; VARANDA, C. & CAPOVILLA, A. G. S. Teste de Competência de Leitura de Palavras e Pseudopalavras: normatização e validação. *PSIC – Revista de Psicologia da Vetor Editora*. v. 7, n. 2, p. 47-69, jul/ dez. 2006.

CAVALIERE, R. A metalinguagem ortográfica na tradição gramatical brasileira. In: SILVA, M. (Org.) *Ortografia da língua portuguesa: história, discurso, representações*. São Paulo: Contexto, 2009. p. 149-164.

CHOMSKY, N. *Knowledge of language: its nature, origin, and use*. Praeger, New York: 1986. 314 p.

CHOMSKY, N. & HALLE, M. *The Sound Pattern of English*. The MIT Press, Cambridge, New York: 1968. 484 p.

- CLEMENTS, G. N. *The geometry of phonological features*. Phonology Yearbook 2: 1985. p. 225-252.
- CLEMENTS, G. N. & HUME, E. V. The internal organization of speech sounds. In: GOLDSMITH, J. A. *The Handbook of Phonological Theory*. Blackwell. 1996.
- COHN, A. C. Laboratory Phonology: past successes and current questions, challenges and goals. In: FOUGERON, C.; KUHNERT, B.; D'IMPERIO, M. & VALEE, N. (eds.). *Laboratory Phonology 10*. Mouton of Gruyter. Berlim & New York, 2010. p. 3- 28.
- COHN, A. C. FOUGERON, C. & HUFFMAN, M. K. *The Oxford Handbook of Laboratory Phonology*. Oxford University Press, New York, 2012.
- COLLISCHONN, G. A sílaba em português. In: BISOL, L. (org.) *Introdução a estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 2010. 5 ed. 286 p.
- CONSONI, F. *O acento lexical como pista para o reconhecimento de palavras: uma análise experimental em palavras segmentadas da língua portuguesa*. 2006. 58 f. Dissertação (Mestrado em Letras). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2006.
- COUTO, H.H. *Fonologia & Fonologia do Português*. Brasília: Thesaurus, 1997.
- CRISTÓFARO-SILVA, T. & OLIVEIRA, M. A. de. *Lateral vocalization in Brazilian Portuguese*. Paper presented at the 3rd UKLVC-University of York, Inglaterra, julho, 2001. p. 1-8.
- CRISTÓFARO-SILVA, T. *Dicionário de Fonética e Fonologia*. São Paulo: Contexto, 2011. 239 p.
- CRISTÓFARO-SILVA, T. *Handout do Curso de Fonologia de Laboratório*. UFMG. 11 a 14 de agosto, 2014.
- CUNHA, C. & CINTRA, L. *Nova gramática do português contemporâneo*. Rio de Janeiro: Lexikon, 2008. 5 ed. 762 p.
- DALGAARD, P. *Introductory Statistic with R*. Springer, 2008, 2 ed. 363 p.
- DEMOLIN, D. The experimental method in Phonology. *Revista da ABRA-LIN*. n. especial 2. Agosto, 2012. p. 31-88.

- ESCUADERO, P. et al. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *Jasa*, v. 126, n. 3, p. 1379-1393, 2009.
- FERREIRA-NETTO, W. O acento na Língua Portuguesa. In: ARAÚJO, G. A. et al. *O acento em português: abordagens fonológicas*. São Paulo: Parábola, 2007. p. 21-36.
- FIELD, A. & HOLE, G. *How to design and report experiments*. SAGE, Londres, 2003. 384 p.
- FLEISCHER, E. Caos/Complexidade na interação humana. In: PAIVA, V. L. M. O. & NASCIMENTO, M. *Sistemas adaptativos complexos: língua(gem) e aprendizagem*. Campinas: Pontes Editores, 2011, 269 p.
- FONSECA, O. Vogais nasais do português: pressupostos e discussão. *Alfa*, São Paulo, 28, p. 101-111, 1984.
- FRANÇA, et al. A neurofisiologia do acesso lexical: palavras em português. *Veredas on-line*. Juiz de Fora, p. 34-39, 2/ 2008. Disponível em: <<http://www.ufjf.br/revistaveredas/files/2009/12/artigo03.pdf>>. Acesso em: 03/06/2014.
- FREELING. *Etiquetas Eagles* (v. 2.0). Disponível em: <<http://www.cis.uni-muenchen.de/schmid/tools/TreeTagger/data/Portuguese-Tagset.html>>. Acesso em: 15/04/2015.
- FRISCH, S. A. Phonotactic Patterns in Lexical Corpora. In: COHEN et al. *The Oxford Handbook of Laboratory Phonology*. Nova Iorque: Oxford University Press, 2012. p. 458- 470.
- FROTA et al. *FrePOP: Frequency Patterns of Phonological Objects in Portuguese: Research and Applications*. Laboratório de Fonética (CLUL), Faculdade de Letras da Universidade de Lisboa. 2010. Disponível em: <<http://frepop.letras.ulisboa.pt/>>. Acesso em 15/10/2015.
- GALVES, C & FARIA, P. (2010). *Tycho Brahe Parsed Corpus of Historical Portuguese*. Disponível em: <<http://www.tycho.iel.unicamp.br/tycho/corpus/en/index.html>>. Acesso em 14/10/2015.
- GARCIA, R. & ZIMMER, M. O papel da frequência lexical e segmental na aquisição das fricativas em crianças de um a três anos: uma perspectiva dinâmica na aquisição do português brasileiro. *Acta Scientiarum. Language and Culture*. Maringá, v. 32, n. 2, p. 279-289, 2010.

- GEERAERTS, D. KRISTIANSEN, G. & PEIRSMAN, Y. Advances in Cognitive Sociolinguistics. In: GEERAERTS, D. & TAYLOR, J. R. *Cognitive Linguistics Research*. New York: Hubert & Co. GmbH & Co. 2010. p. 1-19.
- GONÇALVES, S. C. L. *Banco de dados Iboruna: amostras eletrônicas do português falado no interior paulista*. Disponível em: <<http://www.alip.ibilce.unesp.br/iboruna>>. Acesso em: 08/10/2014.
- GRIES, S. T. *Statistics for Linguistics with R: a practical introduction*. De Gruyter Mouton, Alemanha, 2013. 2 ed. 359 p.
- HAUPT, C. Contribuições da fonologia de uso e da teoria dos exemplares para o estudo da monotongação. *Rev. Est. Ling.*, Belo Horizonte, v. 19, n. 1, p. 167-189, jan./jun. 2011.
- HERMANS, B. & WETZELS, W. L. Productive and unproductive stress patterns in Brazilian Portuguese. *R. Letras & Letras*, Uberlândia – MG, v. 28, n.1, p. 77 – 11, jan./ jun. 2012.
- HUBACK, A. P. da S. Cancelamento do (r) final em nominais: uma abordagem difusionista. *SCRIPTA*, Belo Horizonte, v. 9, n. 18, p. 11-28, 1º sem. 2006.
- HUBACK, A. P. da S. A interferência da frequência em fenômenos linguísticos. *D.E.L.T.A.*, 29:1, p. 79-94, 2013.
- JOHNSON, D. E. Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis. *Language and Linguistics Compass* 3/1, 2009. p. 359-383.
- Johnson, K. Speech perception without speaker normalization: An exemplar model. In: JOHNSON, K. & MULLENNIX, J. W. (eds) *Talker Variability in Speech Processing*. San Diego: Academic Press. pp. 145-165.
- JURAFSKY, D. & MARTIN, J. H. *Speech and Language Processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Prentice Hall, 2008. 2 ed. 1024 p.
- KELLY, M. H., MORRIS, J. & VERREKIA, L. Orthographic cues to lexical stress: Effects on naming and lexical decision. *Memory & Cognition*, v. 26, n. 4, julho/ 1998, p. 822-832.
- KEMMER, S. & BARLOW, M. Introduction: a usage-based conception of language. In: BARLOW, M. & KEMMER, S. *Usage based models of language*.

- Califórnia, CSLI Publications, 2000. p. vii-xxviii.
- KIRCHNER, R. Modeling Exemplar-Based Phonologization. In: COHN, A. C.; FOUGERON, C. & HUFFMAN, M. K. *The Handbook of Laboratory Phonology*. Oxford University Press. New York, 2012. p. 332-344.
- KUHN, T. S. *A estrutura das revoluções científicas*. Perspectiva: São Paulo, 2006. 264 p.
- LABOV, W. *Padrões Sociolinguísticos*. São Paulo: Parábola, 2008. 389 p.
- LADEFOGED, P. & JOHSON, K. *A course in phonetics*. Canadá: Wadsworth Cengage Learning. 2011. 6 ed. 336 p.
- LAGERBERG, R. Variation and frequency in russian word stress. ASEES, Vol. 21, n. 1-2, 2007, p. 165-176.
- LANGACKER, R. W. *Foundations of cognitive grammar*. Stanford: Stanford University, 1987. 1 v.
- LANGACKER, R. W. *Cognitive Grammar: a basic introduction*. Oxford University Press. New York, 2008. 562 p.
- LANGACKER, R. W. *Essentials of Cognitive Grammar*. Oxford University Press: New York, 2013. 269 p.
- LANGACKER, R. A Dynamic Usage-Based Model. In: BARLOW, M. & KEMMER, S. *Usage-based models of language*. CSLI: Califórnia, 1999, p. 1-63.
- LEE, S. *Morfologia e Fonologia Lexical do Português do Brasil*. 1995. 190 f. Tese (Doutorado em Linguística). Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, 1995.
- LIMA, Rocha. *Gramática Normativa da Língua Portuguesa*. Rio de Janeiro: José Olympio, 2002.
- LINGUATECA. *Linguateca*. Disponível em: <<http://www.linguateca.pt/>>. Acesso em: 14/10/2015.
- MANNING, C. D. & SCHÜTZE, H. *Foundations of Statistical Natural Language Processing*. Londres: The MIT Press, 1999. 720 p.
- MARRA, A. V. *Acentuação Gráfica no Português Brasileiro: desafios para a escrita infantil*. 2012. 108 f. Dissertação (Mestrado em Linguística). Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2012.

- MARTINS, M. R. D. *Ouvir Falar: introdução à Fonética do Português*. Caminho, Lisboa, 1988. 187 p.
- MARTINS, M. F. Aspectos da fonologia prosódica Guaraní Mbyá. *Revista Virtual de Estudos de Linguagem - ReVEL*. v. 4, n. 7, agosto de 2006.
- MASSINI-CAGLIARI, G. *Acento e Ritmo*. Contexto, São Paulo, 1992. 95 p.
- MASSINI-CAGLIARI, G. *Do poético ao linguístico no ritmo dos trovadores: três momentos da história do acento*. São Paulo: Cultura Acadêmica Editora, 1999. 207 p.
- MATEUS, M. H. & D'ANDRADE, E. *The Phonology of Portuguese*. Oxford University Press, New York, 2002. 176 p.
- MATEUS, M. H. M., FALÉ, I. & FREITAS, M. J. *Fonética e Fonologia do Português*. Universidade Aberta, Lisboa, 2005. 317 p.
- MATTINGLY, I. G. Reading, Linguistic Awareness, and Language Acquisition. In: DOWING, J. & VALTIN, R. (eds) *Language Awareness and Learning to Read*. Springer: Nova Iorque, 1984. p. 9-25.
- MATTINGLY, I. G. Linguistic Awareness and Orthographic Form. In: FROST, R. & KATZ, L. (Eds.) *Orthography, Phonology, Morphology, and Meaning*. Holanda: Elsevier Science Publishers B. V., 1992. p. 11-26.
- MCENERY, T. & WILSON, A. *Corpus Linguistics: an introduction*. Edinburgh: Edinburgh University Press, 2001. 2 ed. 235 p.
- MEDEIROS, B. R. de. Vogais nasais do português brasileiro: reflexões preliminares de uma revisita. *Revista Letras*, Curitiba, n. 72, p. 165-188, maio/ago. 2007.
- MENDES, R. B. *Projeto SP2010: Amostra da fala paulistana*. 2010. Disponível em <<http://projetosp2010.flch.usp.br>>. Acesso em 30/03/2015.
- MENDES, R. B. & OUSHIRO, L. O paulistano no mapa sociolinguístico brasileiro. *Alfa*, v. 56, n. 3, 2012. p. 973-1001.
- MORAIS, J. et al. Alphabetic literacy and psychological structure. *Letras de Hoje*. Porto Alegre, v. 33, n. 4, p. 61-79, dezembro 1998.
- NATHAN, G. Phonology. In: GEERAERTS, D. & CUYCKENS, H. (eds.) *The Oxford Handbook of Cognitive Linguistics*. Oxford: University Press. 2007. p. 611-631.

- NEY, L. A. G. *Como as crianças utilizam o acento gráfico no processo de aquisição da escrita?* X ANPED SUL, Florianópolis, p.1-22, outubro 2014.
- OHALA, J. J. *There is no interface between phonology and phonetics: a personal view.* Journal of Phonetics, 18. 1990. p. 153-171.
- OHALA, J. J. The relation between Phonetics and Phonology. In: *The Handbook of Phonetic Sciences*. Wiley-Blackwell. West Sussex, 2013. p. 653-677.
- OLIVEIRA, R. A. de. Complexidade: conceitos, origens, afiliações e evoluções. In: PAIVA, V. L. M. O. & NASCIMENTO, M. *Sistemas Adaptativos Complexos: Língua(gem) e Aprendizagem*. Campinas: Pontes, 2011, p. 13-34.
- OLIVEIRA, M. A. de & CRISTÓFARO-SILVA, T. Variação do "r" pós-consonantal no português brasileiro: um caso de mudança fonotática. *Letras de Hoje*. Porto Alegre, v. 37, n.1, p. 25-47, março. 2002.
- OUSHIRO, L. *Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*. 2015. 372 f. Tese (Doutorado em Letras). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2015.
- PAIVA, V. L. M. O. & NASCIMENTO, M. (orgs.) *Sistemas adaptativos complexos: língua(gem) e aprendizagem*. São Paulo: Pontes, 2011. 270 p.
- PIERREHUMBERT, J. B. Exemplar dynamics: word frequency, lenition and contrast. In: BYBEE, J. & HOPPER, P. (eds.). *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins, 2001. p. 137-157.
- PIERREHUMBERT, J. B. & BECKMAN, M. E. & LADD, D. Robert. Conceptual foundations of phonology as a laboratory science (reprint). In: COHN, Abigail C. FOUGERON, Cécile & HUFFMAN, Marie K. *The Handbook of Laboratory Phonology*. Oxford University Press. New York, 2012. 17-39.
- PIPERSKI, A. Ch. & KUKHTO, A. V. Intra-speaker stress variation in russian: a corpus-driven study of russian poetry. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2016"*. Moscow, junho, 2016.
- Python Team. *Python Software Foundation*. The Python Software Foun-

dation is a non-profit corporation. Versão 2.7.1. 2015. Disponível em <<https://www.python.org/downloads/release/python-2710rc1/>>.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. 2015. Disponível em <<http://www.R-project.org>>.

RASO, T. & MELLO, H. (Org.) *C-ORAL-BRASIL: Corpus de referência do português brasileiro falado informal*. Belo Horizonte: Editora UFMG, 2012. 332 p.

RODRIGUES, J. C. & SALLES, J. F. Tarefa de escrita de palavras/pseudopalavras para adultos: abordagem da neuropsicologia cognitiva. *Letras de Hoje*, Porto Alegre, v. 48, n. 1, p. 50-58, jan./ mar. 2013.

SARDINHA, T. B. (2000a) Linguística de Corpus: Histórico e Problemática. *DELTA*, v. 16, n. 2, 2000. p. 323-67.

SARDINHA, T. B. (2000b) *O que é um corpus representativo?* Direct Papers 44. Disponível em: <<http://www.direct.f2s.com>>. Acesso em: 02/09/2014.

SAUSSURE, Ferdinand. *Curso de Linguística Geral*. Cultrix, São Paulo: 1972. 4 ed. 280 p.

SCHILLER, N. O. Experimental methods and designs to investigate phonological encoding of spoken language. In: COHN, A. C. FOUGERON, C. & HUFFMAN, M. K. *The Oxford Handbook of Laboratory Phonology*. Oxford University Press, New York, 2012. p. 528-538.

SCHMID, H. *Probabilistic Part-of-Speech Tagging Using Decision Trees*. Proceedings of International Conference on New Methods in Language Processing, Manchester, UK. 1994. p. 1-9.

SCHMID, H. *Improvements in part-of-speech tagging with a application to German*. Proceedings of the ACL SIGDAT-Workshop. Dublin, Ireland. 1995. p. 1-9.

SCHMID, H. *TreeTagger: a language independent part-of-speech tagger*. Disponível em: <<http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/>>. Acesso em: 15/04/2015.

SEIDENBERG, M. S. & TANENHAUS, M. K. Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning*

and Memory. v. 5, n. 6, p. 546-554, 1979.

SELKIRK, E. O. The syllable. In: HULST, H; SMITH, V. D. *The structure of phonological representations* (part. II). Foris, Dordrecht, p. 337-383. 1982.

SHADEMAN, S. *Is Phonotactic Knowledge Grammatical Knowledge?*. Proceedings of the 25th West Coast Conference on Formal Linguistics. Somerville, MA: Cascadilla Proceedings Project. 2006. p. 371-379.

SIMIONI, T. *A alternância entre ditongo crescente e hiato em português: uma análise otimalista*. 2005. 155 f. Dissertação (Mestrado em Letras). Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

SOUZA, L. C. S. & PACHECO, V. Uma análise acústica das vogais orais, nasais e nasalização no dialeto de Vitória da Coquista, Bahia. *SIGNUM: Estud. Ling.*, Londrina, n. 15/2, p. 401-431, dez. 2012.

STUART, D. *On the internal organization of syllable constituents*. University of Arizona Linguistics Circle. Disponível em: <<http://hdl.handle.net/10150/226099>>, p. 3-22, 1982.

TEYSSIER, P. *História da Língua Portuguesa*. Trad. Sob a direção de Celso Cunha. 3 ed. São Paulo: Martins Fontes, 2007. 142 p.

VELOSO, J. M. P. da S. e A. *Da influência do conhecimento ortográfico sobre o conhecimento fonológico: estudo longitudinal de um grupo de crianças falantes nativas do português europeu*. Tese (Doutorado em Linguística). Faculdade de Letras da Universidade do Porto, Porto, 2003. f. 505.

VIARO, M. E. & GUIMARÃES-FILHO, Z. O. Análise quantitativa da frequência dos fonemas e estruturas silábicas portuguesas. *Estudos Lingüísticos XXXVI(1)*, janeiro-abril, 2007. p. 27 – 36.

VITEVITCH, M. S. & LUCE, P. A. *Probabilistic Phonotactics and Neighborhood Activation in Spoken Word Recognition*. Journal of Memory and Language, 1999, 40, p. 374-408.

VITEVITCH, M. S. et al. *Phonotactics and Syllable Stress: implications for the processing of spoken nonsense words*. Language and Speech, 1997, 40(1), p. 47-62.

VITEVITCH, M. S. et al. *Phonotactics, Neighborhood Activation, and Lexical Access for Spoken Words*. Brain and Language, 68, 1999, p. 306-311.

Anexo A: Resultados do metaexperimento

Tabela 1 Resultados gerais de atribuição acentual por *condição experimental* e *pseudopalavra*.

Pseudopalavras	Áudio			Imagem			Ortografia		
	Prop	P	O	Prop	P	O	Prop	P	O
Casemi	0	1	8	0	9	0	0	0	9
Citaba	0	6	3	1	8	0	0	9	0
Fuchedo	0	9	0	0	9	0	0	9	0
Lepaque	0	7	2	0	9	0	0	9	0
Gazonar	0	0	9	0	0	9	0	0	9
Piredir	0	0	7	0	0	9	0	0	9
Rodipam	0	0	9	0	0	9	0	0	9
Vunhatil	0	2	7	0	3	6	0	1	8
Ducirca	0	5	4	0	9	0	0	9	0
Jofelta	0	6	3	0	9	0	0	9	0
Malanto	0	9	0	0	9	0	0	9	0
Zubesma	0	7	2	0	9	0	0	9	0
Bitosper	0	3	6	0	3	6	0	3	6
Chidesvor	0	2	7	0	2	7	0	3	6
Nefurdam	0	0	9	0	1	8	0	0	9
Tubandal	0	0	9	0	0	9	0	0	9
Cindoto	0	1	0	6	3	0	0	8	0
Dercali	0	7	2	0	9	0	0	7	2
Gasbochu	0	4	5	0	9	0	0	5	4
Tospurru	0	6	3	0	3	6	0	6	3

Anexo B:

Resultados estatísticos do metaexperimento

Figura 1 Resultados do modelo de regressão logística binomial do metaexperimento.

```
*****
Family:  c("BI", "Binomial")

Call:  gamlss(formula = TONICIDADE ~ PSEUDOPALAVRA + REPETICAO + ESCOLARIDADE +
      EstrSilabaTon + FreqEstrPalavra + PosicaoSilPesada + ESTIMULO +
      random(INFORMANTE), family = BI, data = dados)

Fitting method: RS()

-----
Mu link function:  logit
Mu Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)    -0.2337    0.8248   -0.283 0.777043
PSEUDOPALAVRACasemi    -1.3821    0.8654   -1.597 0.110879
PSEUDOPALAVRACHidesvor    -0.4200    0.6477   -0.648 0.516985
PSEUDOPALAVRACindoto    13.7105   410.0562    0.033 0.973341
PSEUDOPALAVRACitaba     1.7686    1.0036    1.762 0.078642 .
PSEUDOPALAVRADercali     1.5989    0.9046    1.768 0.077756 .
PSEUDOPALAVRADucirca     2.8452    0.7392    3.849 0.000134 ***
PSEUDOPALAVRAFuchedo    14.3125   255.7324    0.056 0.955391
PSEUDOPALAVRAGasbochu     0.1051    0.8778    0.120 0.904720
PSEUDOPALAVRAGazonar   -13.8221   265.7370   -0.052 0.958539
PSEUDOPALAVRAJofelta     3.2791    0.8339    3.932 9.63e-05 ***
PSEUDOPALAVRALepaque     2.3024    1.0848    2.123 0.034296 *
PSEUDOPALAVRAMalanto    16.0457   255.7317    0.063 0.949996
PSEUDOPALAVRANefurdam    -2.6007    1.1266   -2.308 0.021393 *
PSEUDOPALAVRAPiredir    -2.6181    1.1257   -2.326 0.020439 *
PSEUDOPALAVRARodipam   -13.8221   265.7370   -0.052 0.958539
PSEUDOPALAVRATospurru    -0.5047    0.8636   -0.584 0.559219
PSEUDOPALAVRATubandal   -13.8221   265.7370   -0.052 0.958539
PSEUDOPALAVRAVunhatil   -1.0970    0.7295   -1.504 0.133288
PSEUDOPALAVRAZubesma     4.0356    0.9222    4.376 1.48e-05 ***
REPETICA02    -0.2596    0.3705   -0.701 0.483728
REPETICA03    -0.6530    0.3737   -1.747 0.081195 .
ESCOLARIDADEsuperior     0.5063    0.3246    1.560 0.119474
EstrSilabaTonCVC    -1.7332    0.6332   -2.737 0.006421 **
ESTIMULOimagem     1.8366    0.3994    4.598 5.43e-06 ***
ESTIMULOortografia     0.9111    0.3758    2.424 0.015692 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Anexo C:

Pré-teste de controle de tipos

Este anexo apresenta uma breve síntese das motivações e dos resultados obtidos em um pré-teste que objetivou avaliar o método experimental e levantar hipóteses a respeito de uma abordagem para o acento no PB a partir de tipos silábicos.

Pré-teste

Entre as diversas variáveis que parecem interagir e influir no comportamento acentual, definiu-se como ponto de investigação para este pré-teste o controle dos tipos silábicos. Isso porque a emergência consistente de proparoxítonas em um contexto silábico específico, *sin-*, a produção oscilante entre *vun~~h~~atil/vunhatil* e *chides~~v~~or/chidesvor* e as ocorrências ambíguas nos segmentos finais *i* e *u*, como em *casemi*, *dercali*, *tospurru* e *gasbochu*, no metaexperimento, impeliram questões a respeito da frequência dos tipos silábicos na gramática fonológica do falante, a qual pode se comportar como preditora do acento em português.

Os casos elencados acima foram escolhidos como sílaba/segmento alvo para compor um teste-piloto, o qual objetivou aprimorar o desenho experimental da condição imagem realizado no metaexperimento, bem como investigar em que medida os tipos silábicos podem se comportar como atratores para o acento no PB. Para que a investigação pudesse ser realizada, criaram-se 15 pseudopalavras, 3 para cada tipo silábico final, como pode ser consultado na Tabela 2, os quais foram apresentados a 4 falantes nativos do português.

Tabela 2 Pseudopalavras por sílaba/segmento alvo.

	[sĩ-]	[-i]	[-u]	[-iw]	[-vor]
Pseudopalavras	[sĩboko]	[kẽpiri]	[metobu]	[kavuniw]	[kalisvor]
	[sĩzuta]	[pukari]	[perfutu]	[fematfiw]	[dekarvor]
	[sĩfuzo]	[tadevi]	[zapulu]	[sibetfiw]	[sibervor]

Os resultados gerais corresponderam a 1 (0,6%) produção proparoxítona, 96 (53,6%) paroxítonas e 82 (45,8%) oxítonas, com uma diferença significativa de $\chi^2 = 88.2$ (2), $p < 0.001$, conforme demonstra a Tabela 3.

Tabela 3 Atribuição acentual no pré-teste.

Proparoxítona	Paroxítona	Oxítona
1 (0,6%)	96 (53,6%)	82 (45,8%)

$$\chi^2 = 88.2$$
 (2), $p < 0.001$

De maneira geral, como pode ser vista na Tabela 4, observou-se preferência por acento paroxítono em estruturas de palavras com a sílaba final leve, à medida que há preferência pelo acento oxítono quando a sílaba final é ramificada. Contudo, em ambos os casos, a quantidade de produções oxítonas e paroxítonas, respectivamente, é notável, fato que evidencia que a estrutura de palavra não é o único elemento relevante para a atribuição acentual.

Ao observar os tipos silábicos, percebeu-se que o acento tende a variar mais em pseudopalavras em que a estrutura de palavra favorece o acento paroxítono e as informações segmentais favorecem o acento oxítono, o que indicaria uma competição entre essas variáveis, como em [-i] e [-u]. Além disso, constatou-se que, seguindo a predição de que a frequência dos tipos silábicos tem impacto nas representações fonológicas, não houve produções paroxítonas com [niw], somente com [tʃiw], o qual é encontrado na língua em *versátil*, *erétil*, *réptil*. Em contrapartida, atestaram-se produções paroxítonas inesperadas em [kalisvor] e [dekarvor], mas não em [sibervor]. Esses resultados incitam questionamentos a respeito das motivações que conduzem à produção ora do acento paroxítono, ora do acento oxítono, mostrando que fatores como a frequência dos tipos silábicos e, até mesmo, os processos analógicos podem

Tabela 4 Atribuição acentual por pseudopalavra no pré-teste.

Pseudopalavra	Proparoxítona	Paroxítona	Oxítona
[sĩboko]	0	12	0
[sĩfuzo]	0	12	0
[sĩʒuta]	0	12	0
[kẽpiri]	0	9	3
[pukari]	0	8	4
[tadevi]	0	9	3
[metobu]	0	6	6
[perfutu]	0	12	0
[zapulu]	1	4	7
[kavuniw]	0	0	12
[fematʃiw]	0	5	7
[sibetʃiw]	0	2	10
[kalisvor]	0	3	9
[dekarvor]	0	2	10
[sibervor]	0	0	11

ser fatores que motivam a alternância acentual.

Considerações Finais

Os mecanismos que guiam o acento em pseudopalavras ainda não são claros o suficiente para se tecer qualquer conclusão a respeito dessa ou de outra variável como primordial para a atribuição acentual. Até então, pode-se apenas concluir que os resultados deste teste são fortes o bastante para se assegurar que o acento em português não pode ser explicado por meio de casos categóricos, que o acento proparoxítono consiste num padrão da língua e, principalmente, que, a depender do item lexical, competições acentuais podem emergir, viabilizando mais de uma produção acentual.

Anexo D:

Gráfico do modelo de regressão logística no Grupo de Análise 1

Figura 2 Gráfico de adequação do modelo de regressão logística no Grupo de Análise 1

Anexo E:

Gráfico de adequação do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 1

Figura 3 Erros do modelo de regressão logística binomial com efeitos mistos em $[niw]$, $[V_{oral.t}f_{iw}]$ e $[V_{nasal.t}f_{iw}]$.

Anexo F:

Resultados do modelo no Grupo de Análise 1

Figura 4 Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 1 no modelo de [grupo de análise + sílaba final + repetição].

```
Call:  gamlss(formula = TONICIDADE ~ ALVO + SILABA.FINAL + REPETICAO +
             random(INFORMANTE), family = BI, data = na.omit(dados.geral.noPROP))

Fitting method: RS()

-----
Mu link function:  logit
Mu Coefficients:

```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.12985	0.30915	6.889	6.88e-12	***
ALVOntil	-2.43928	0.31838	-7.662	2.50e-14	***
ALVOtil	-1.78543	0.33174	-5.382	7.97e-08	***
SILABA.FINALantil	0.74086	0.29058	2.550	0.0108	*
SILABA.FINALatil	-0.46980	0.29014	-1.619	0.1055	
SILABA.FINALenil	0.59227	0.36020	1.644	0.1002	
SILABA.FINALentil	-0.39048	0.26257	-1.487	0.1371	
SILABA.FINALetil	-0.29570	0.29328	-1.008	0.3134	
SILABA.FINALunil	0.19357	0.36141	0.536	0.5923	
REPETICAO	0.41754	0.08768	4.762	2.01e-06	***

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
NOTE: Additive smoothing terms exist in the formulas:
i) Std. Error for smoothers are for the linear effect only.
ii) Std. Error for the linear terms may not be reliable.

-----
No. of observations in the fit: 2876
Degrees of Freedom for the fit: 44.39317
Residual Deg. of Freedom: 2831.607
at cycle: 4

Global Deviance: 1263.129
AIC: 1351.915
SBC: 1616.683
*****

```

Figura 5 Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 1 no modelo de [pseudopalavra + repetição].

```
Call:
gamlss(formula = TONICIDADE ~ PSEUDOPALAVRA + REPETICAO + random(INFORMANTE),
        family = BI, data = na.omit(dados.geral.noPROP))

Fitting method: RS()

-----
Mu link function: logit
Mu Coefficients:

```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.12285	0.38668	-0.318	0.750744
PSEUDOPALAVRAbikuntíl	-0.17721	0.49057	-0.361	0.717954
PSEUDOPALAVRAbudantíl	1.90411	0.56812	3.352	0.000814 ***
PSEUDOPALAVRAcabentíl	0.52236	0.51525	1.014	0.310763
PSEUDOPALAVRAcobatíl	-0.30328	0.48520	-0.625	0.531974
PSEUDOPALAVRAcorruntíl	3.36780	0.57169	5.891	4.30e-09 ***
PSEUDOPALAVRADerruntíl	0.52236	0.51525	1.014	0.310763
PSEUDOPALAVRAfecatíl	0.24687	0.50277	0.491	0.623450
PSEUDOPALAVRAgoletíl	-0.24083	0.48212	-0.500	0.617454
PSEUDOPALAVRALifentíl	2.55203	0.57462	4.441	9.29e-06 ***
PSEUDOPALAVRALopeníl	3.75095	0.56767	6.608	4.65e-11 ***
PSEUDOPALAVRAMosantíl	0.92840	0.54032	1.718	0.085860 .
PSEUDOPALAVRAMusantíl	3.35780	0.57345	5.855	5.31e-09 ***
PSEUDOPALAVRAnatutíl	0.99963	0.53832	1.857	0.063425 .
PSEUDOPALAVRApetuntíl	2.52976	0.57829	4.375	1.26e-05 ***
PSEUDOPALAVRApolentíl	-1.31139	0.45271	-2.897	0.003799 **
PSEUDOPALAVRAREbuntíl	-0.55969	0.47081	-1.189	0.234630
PSEUDOPALAVRASilantíl	1.72085	0.56190	3.063	0.002215 **
PSEUDOPALAVRASimatíl	-0.12323	0.48635	-0.253	0.799999
PSEUDOPALAVRASipantíl	0.38676	0.50815	0.761	0.446653
PSEUDOPALAVRASupentíl	0.51893	0.51556	1.007	0.314247
PSEUDOPALAVRAtajentíl	2.50713	0.57480	4.362	1.34e-05 ***
PSEUDOPALAVRAtaventíl	-0.92390	0.46171	-2.001	0.045486 *
PSEUDOPALAVRaticuntíl	1.65319	0.57018	2.899	0.003768 **
PSEUDOPALAVRatijutíl	-0.15670	0.48845	-0.321	0.748372
PSEUDOPALAVRAvenantíl	0.25642	0.51325	0.500	0.617395
PSEUDOPALAVRAvenetíl	0.11601	0.49676	0.234	0.815365
REPETICAO	0.43449	0.08952	4.854	1.28e-06 ***

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
-----
```

Anexo G:

Resultados do modelo no

Grupo de Análise 2

Figura 6 Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 2 no modelo de [grupo de análise + sílaba final + repetição].

```
Call:  gamlss(formula = TONICIDADE ~ ALVO + SILABA.FINAL + REPETICAO +
             random(INFORMANTE), family = BI, data = na.omit(dados.geral.noPROP))

Fitting method: RS()

-----
Mu link function:  logit
Mu Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.34938    0.37173   9.010  <2e-16 ***
ALV0or         -0.75627    0.36380  -2.079   0.0378 *
SILABA.FINALgor -0.33162    0.34385  -0.964   0.3350
SILABA.FINALtar  0.33391    0.37493   0.891   0.3733
SILABA.FINALtor -0.09042    0.36095  -0.251   0.8022
SILABA.FINALvar  0.66308    0.38062   1.742   0.0817 .
REPETICAO       0.11719    0.12915   0.907   0.3643
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

-----
NOTE: Additive smoothing terms exist in the formulas:
 i) Std. Error for smoothers are for the linear effect only.
ii) Std. Error for the linear terms may not be reliable.
-----
No. of observations in the fit: 1859
Degrees of Freedom for the fit: 38.33509
Residual Deg. of Freedom: 1820.665
at cycle: 5

Global Deviance: 586.5608
AIC: 663.231
SBC: 875.1395
*****
```


Figura 7 Resultados do modelo de regressão logística binomial com efeitos mistos no Grupo de Análise 2 no modelo de [pseudopalavra + repetição].

```
Call:
gamlss(formula = TONICIDADE ~ PSEUDOPALAVRA + REPETICAO + random(INFORMANTE),
family = BI, data = na.omit(dados.geral.noPROP))

Fitting method: RS()
```

```
Mu link function: logit
Mu Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	2.49777	0.51251	4.874	1.19e-06	***
PSEUDOPALAVRAcamitor	-0.07905	0.65597	-0.121	0.90409	
PSEUDOPALAVRAcapigor	0.03016	0.60827	0.050	0.96046	
PSEUDOPALAVRAdefivor	0.59868	0.62429	0.959	0.33769	
PSEUDOPALAVRadumitar	1.64437	0.63832	2.576	0.01007	*
PSEUDOPALAVRAfelitor	0.28868	0.63209	0.457	0.64794	
PSEUDOPALAVRagabitar	0.41853	0.62338	0.671	0.50206	
PSEUDOPALAVRALitivor	-0.17288	0.60501	-0.286	0.77510	
PSEUDOPALAVRalogivar	1.51563	0.65165	2.326	0.02014	*
PSEUDOPALAVRALufigar	1.32528	0.65138	2.035	0.04204	*
PSEUDOPALAVRAmedigor	-0.69292	0.59279	-1.169	0.24259	
PSEUDOPALAVRAnovitar	1.83786	0.65798	2.793	0.00527	**
PSEUDOPALAVRApessivar	1.21802	0.63457	1.919	0.05508	.
PSEUDOPALAVRApossigor	0.03966	0.60990	0.065	0.94815	
PSEUDOPALAVRArapivar	2.23249	0.67822	3.292	0.00102	**
PSEUDOPALAVRAsetigar	1.51949	0.64632	2.351	0.01883	*
PSEUDOPALAVRatanivor	-0.02109	0.63278	-0.033	0.97341	
PSEUDOPALAVRAzibitor	-0.09545	0.61895	-0.154	0.87746	
REPETICAO	0.13066	0.13152	0.993	0.32064	

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```