

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
Faculdade de Letras  
Programa de Pós-graduação em Estudos Linguísticos

Vitor Antônio Morais Prado

## **O acento lexical e seu impacto nos constituintes silábicos**

Belo Horizonte

2025

Vitor Antônio Morais Prado

## **O acento lexical e seu impacto nos constituintes silábicos**

Dissertação apresentada ao Programa de Pós-graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de Mestre em Linguística.

Orientadora: Profa. Dra. Maria Mendes Cantoni

Belo Horizonte

2025

P896a Prado, Vitor Antônio Morais.  
O acento lexical e seu impacto nos constituintes silábicos [manuscrito] / Vitor Antônio Morais Prado. – 2025.  
1 recurso online (140 f. : il., tabs., grafs. (algumas color.)) : pdf.  
Orientadora: Maria Mendes Cantoni.  
Área de concentração: Linguística Teórica e Descritiva.  
Linha de Pesquisa: Estudos Formais em Língua.  
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.  
Bibliografia: f. 102-113.  
Anexos: f. 114-119.  
Apêndices: f. 120-140.  
Exigências do sistema: Adobe Acrobat Reader.

1. Língua portuguesa – Fonologia – Teses. 2. Língua portuguesa – Ortografia e silabação – Teses. 3. Língua portuguesa – Acentos e acentuação – Teses. I. Cantoni, Maria Mendes. II. Universidade Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD : 469.15



UNIVERSIDADE FEDERAL DE MINAS GERAIS

## FOLHA DE APROVAÇÃO

**O acento lexical e seu impacto nos constituintes silábicos**

**VITOR ANTÔNIO MORAIS PRADO**

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGÜÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGÜÍSTICOS, área de concentração LINGÜÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Estudos Formais de Língua.

Aprovada em 19 de março de 2025, pela banca constituída pelos membros:

Prof(a). Maria Mendes Cantoni - Orientadora

UFMG

Prof(a). Magnun Rochel Madruga

UFMG

Prof(a). Aline de Lima Benevides

FATEC

Belo Horizonte, 19 de março de 2025.



Documento assinado eletronicamente por **Maria Mendes Cantoni, Professora do Magistério Superior**, em 20/03/2025, às 16:03, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Aline de Lima Benevides, Usuário Externo**, em 21/03/2025, às 17:04, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



Documento assinado eletronicamente por **Magnun Rochel Madruga, Professor do Magistério Superior**, em 24/03/2025, às 23:47, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).

---



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **4056229** e o código CRC **1CC78E79**.

---

*Aos meus professores e professoras, e aos que os formaram.*

# Agradecimentos

O presente trabalho não existiria sem o auxílio de diversas pessoas, além do próprio autor.

À minha orientadora, Maria Cantoni. Sou mais feliz por ter tido a oportunidade de ter a sua orientação e amizade ao longo dos anos de mestrado. A presente dissertação muito se beneficiou de estar sob o olhar de uma linguista e fonóloga brilhante. Obrigado por me levar além do que eu jamais poderia alcançar sozinho e, assim, ampliar meus horizontes.

Aos membros da banca examinadora. Sou grato desde já pelas contribuições que serão trazidas por novos olhares. Poder ter o trabalho examinado e avaliado por outros fonólogos, que admiro, é um privilégio que será aproveitado e muito apreciado.

Ao professor Magnun Rochel Madruga. Obrigado pelos diversos comentários e contribuições a este trabalho ao longo do meu mestrado. Muitas das habilidades de programação e de análise estatística que eu desenvolvi nestes dois anos, eu devo a você.

À professora Luciana Lucente. Sou grato pelos seus comentários e explicações sobre fonética e padrões entoacionais do português brasileiro, que me fizeram pensar “acima” dos segmentos. Muito obrigado.

Ao professor Rui Rothe-Neves. Agradeço por ter despertado em mim uma profunda curiosidade sobre o acento de palavra no português, através da excelente orientação na Iniciação Científica sobre choque acentual. Sou imensamente grato pelos ensinamentos em Fonética e por ter me acolhido desde os primeiros passos na graduação, quando eu ainda era apenas um calouro.

Aos professores das disciplinas obrigatórias que cursei ao longo do mestrado, que me formaram para ser verdadeiramente um mestre em Linguística. Obrigado, Janayna Carvalho e Fábio Bonfim, por me apresentarem o gerativismo e fazerem, incrivelmente, eu me apaixonar por ele. Obrigado, César Nardelli, pelas brilhantes e insubstituíveis aulas de Geossociolinguística. Obrigado, Aroldo de Andrade, pelas excelentes aulas sobre contato linguístico.

À professora Márcia Regina Jaschke Machado. Agradeço imensamente pelas aulas de Literatura Brasileira durante a minha graduação, especialmente sobre o Movimento Modernista Paulista e tantos outros temas que me transformaram profundamente. Suas aulas me deram as ferramentas para escapar da linguística sempre que precisei, encontrando refúgio nas artes de vanguarda e experimentais. Durante as pausas desta dissertação, foram os filmes de Truffaut e os textos de Luís Aranha e Guilherme de Almeida – que eu jamais conheceria sem você – que, de alguma forma, me mantiveram firme no caminho que escolhi dentro da fonética e fonologia. Dos modernistas é a culpa se a sintaxe do texto que vem a seguir por vezes for demais embaralhada. Serei para sempre grato por isso.

Aos colegas e equipe do Laboratório de Fonologia da UFMG, principalmente à fundadora, Professora Thaís Cristófaros Silva, e aos coordenadores, Maria Cantoni e Magnun Madruga. Foi um privilégio poder me abrigar em um ambiente tão adequado para as tarefas acadêmicas, com equipamentos fundamentais para que o presente trabalho se realizasse. Estivesse o autor sujeito ao estado comum e atual de infraestrutura das instalações da Universidade Pública brasileira, especificamente das faculdades de ciências humanas e linguagens, a qualidade desta dissertação estaria comprometida. Para além disso, o laboratório foi espaço de cultivo de novas amizades. Agradeço ao Ariel pelo apoio e companhia durante as pausas dos estudos, e aos demais colegas, que fizeram da academia um lugar menos entediante: Daniel, Flora, os dois Gabriels, Julia, Luíza, Maria Elisa, Max, Matheus, Melissa, Rosivaldo, Tim e Jennyffer.

Aos meus pais e à minha irmã. Obrigado por me apoiarem incondicionalmente, mesmo sem entenderem exatamente o que estava acontecendo ao longo desses dois anos. O carinho e paciência inabalável em mim me sustentaram de certa forma em cada etapa deste caminho.

Aos meus amigos. Ao Jorlan, por todo o carinho e pelo constante não-proposital ato de me lembrar do que eu também amo nas ciências da linguagem: a literatura e o ensino de língua materna. À Arielle, por tanta companhia e amor, durante todos esses anos, apesar de tanta teoria linguística inundando nossos encontros. À Rayna, pelas risadas e conselhos, e por ter sido um dos maiores presentes que a Universidade me deu. À Aline, pela companhia durante o mestrado, que foi essencial para não desistir em momentos difíceis e para me fazer ter coragem de enfrentar teorias além das fonológicas. Ao João Ricardo, pela incrível amizade a um fuso-horário de distância, que fez com que eu pudesse sair um pouco de Belo Horizonte quando quisesse. Agradeço também ao Raphael, à Victória, à Laura, e a tantos outros queridos amigos.

À Universidade Federal de Minas Gerais, à Faculdade de Letras e ao Programa de Pós-Graduação em Estudos Linguísticos, pelo ensino público, que tanto resiste, garantindo que o conhecimento e a ciência sejam direitos, não privilégios.

À CAPES, pelo apoio financeiro, sem o qual esta dissertação literalmente não existiria.

Por fim, a todos que, mesmo sem serem nomeados, contribuíram para que este trabalho se tornasse realidade, meu sincero agradecimento.

*“From the ears grew two tongues  
And I sang for people passing  
A strange song  
Told them stories without moving my lips  
Mouth half-open, still  
They assumed the words came from themselves  
These unfamiliar thoughts.”  
(Jenny Hval, 2011, *Blood flight*, do álbum *Viscera*)*

# Resumo

Este trabalho analisa o efeito de complexidade silábica e de posição segmental na sílaba nos padrões de duração associados ao acento de palavra no português brasileiro (PB). Estudos sobre acento analisam predominantemente sílabas CV (MOUSIKOU; STRYCHARCZUK; RASTLE, 2024; VOGEL; ATHANASOPOULOU; PINKUS, 2016; CANTONI, 2013) e focam no conteúdo vocálico da sílaba, i.e. em seu núcleo. Argumenta-se que a duração das sílabas acentuadas em relação às átonas depende do molde silábico, assim como a duração dos segmentos acentuados depende de sua posição dentro da sílaba, conforme alguns entendimentos de realinhamento temporal da Fonologia Articulatória (FA) (BROWMAN; GOLDSTEIN, 1986), a saber, o efeito Centro-C. Para a FA, o início de articulação gestual em ataques complexos – mas não em codas – segue um alinhamento em torno de um centro, fato que altera o início de articulação dos gestos (HALL, 2010). Relaciona-se a propriedade de deslocamento temporal à alteração duracional presente em gestos sob o efeito (ou não) do acento de palavra. Considerar segmentos não-nucleares permite averiguar a interação entre o alinhamento em Centro-C e o acento de palavra. Foi feito um experimento com 14 falantes do PB. A tarefa consistiu na leitura de 282 frases com palavras-alvo contendo sílabas de moldes C(C)V(C), tanto tônicas, quanto postônicas, e iniciadas por consoantes de diferentes características articulatórias. A vogal-núcleo foi mantida [a] ou [ɐ]. Para tornar mais clara a interpretação do problema, a variável dependente, duração, foi transformada em taxa de duração entre elementos postônicos e tônicos correspondentes. Os resultados apontam que sílabas tônicas CV são 81,4% maiores que seu pares átonos. Os outros moldes apresentaram taxas menores: 58,6% (CVC); 69% (CCV); e 52,9% (CCVC). O núcleo da sílaba não foi sensível a diferentes moldes. As consoantes tônicas em ataque simples são, em média, 27,4% maiores do que seus pares átonos, porém sensível a efeitos de molde silábico, modo de articulação e vozeamento (com interações entre efeitos), sendo as mais alongadas aquelas fricativas, desvozeadas ou em sílabas CV. O tepe, embora tenha sofrido um leve alongamento em posição acentuada (+13%), não teve o molde silábico como um efeito relevante em sua duração. Por fim, em posição de coda, o /s/ aponta estar sensível à tonicidade da sílaba seguinte, e não de sua própria. Enquanto ainda de acordo com o esperado pelo efeito Centro-C, os resultados para a coda proporcionam uma discussão cara aos estudos de encontros-sC (*sC-clusters*) e de considerações sobre a existência de outras unidades rítmicas, como o IPCG (BARBOSA, 1996). Esses resultados corroboram a hipótese de que a complexidade do molde silábico e a posição do segmento interagem com a manifestação do acento de palavra.

**Palavras-chave:** acento, sílaba, duração, Fonologia Articulatória.

# Abstract

This study analyzes the effects of syllabic complexity and segmental position within the syllable on the duration patterns associated with word stress in Brazilian Portuguese (BP). Research on stress predominantly focuses on CV syllables (MOUSIKOU; STRYCHARCZUK; RASTLE, 2024; VOGEL; ATHANASOPOULOU; PINKUS, 2016; CANTONI, 2013) and the vocalic content of the syllable, i.e., its nucleus. It is argued that the duration of stressed syllables relative to unstressed ones depends on the syllabic template, just as the duration of stressed segments depends on their position within the syllable, in line with certain temporal realignment principles of Articulatory Phonology (AP) (BROWMAN; GOLDSTEIN, 1986), particularly the c-center effect. According to AP, the onset of gestural articulation in complex onsets—but not in codas—aligns around a central point, altering the initiation of articulation (HALL, 2010). The temporal displacement property is thus linked to the durational variation observed in gestures under the influence (or not) of word stress. Considering non-nuclear segments allows for an investigation of the interaction between c-center alignment and word stress. An experiment was conducted with 14 BP speakers. The task involved reading 282 sentences containing target words with syllables of the C(C)V(C) structure, both stressed and post-tonic, and beginning with consonants of different articulatory properties. The nuclear vowel was either [a] or [ɐ]. To clarify the interpretation of the problem, the dependent variable, duration, was transformed into a ratio comparing post-tonic and stressed elements. Results indicate that stressed CV syllables are 81.4% longer than their unstressed counterparts. Other syllabic structures exhibited lower ratios: 58.6% (CVC), 69% (CCV), and 52.9% (CCVC). The syllabic nucleus was not sensitive to different templates. Stressed consonants in simple onsets were, on average, 27.4% longer than their unstressed counterparts, but their duration was influenced by syllabic structure, manner of articulation, and voicing (with interaction effects). The longest segments were fricatives, voiceless, or in CV syllables. The tap, while slightly lengthened in stressed position (+13%), was not significantly affected by syllabic structure. Finally, in coda position, /s/ appeared to be sensitive to the stress of the following syllable rather than its own. While consistent with the expected c-center effect, the coda results contribute to discussions on sC-clusters and the existence of additional rhythmic units, such as IPCG (BARBOSA, 1996). These findings support the hypothesis that syllabic complexity and segmental position interact with the manifestation of word stress.

**Keywords:** stress, syllable, duration, Articulatory Phonology.

# Lista de ilustrações

Figura 1 – Proposta de análise de alongamento da sílaba como um todo em quatro moldes silábicos. . . . .	27
Figura 2 – Proposta de análise de alongamento de cada segmento da sílaba em quatro moldes silábicos. . . . .	27
Figura 3 – Representação da estrutura silábica de <paz> de acordo com um modelo de estrutura linear ou plana. . . . .	31
Figura 4 – Representação da estrutura silábica de <paz> de acordo com um modelo de ramificação à direita. . . . .	32
Figura 5 – Representação da estrutura silábica de <paz> de acordo com um modelo de ramificação à esquerda. . . . .	32
Figura 6 – Representação da estrutura silábica de <paz> de acordo com a Fonologia de Governo. . . . .	32
Figura 7 – Representação da estrutura silábica de <paz> de acordo com um modelo moraico. . . . .	33
Figura 8 – Representação da estrutura silábica de <paz> de acordo com um modelo ternário. . . . .	33
Figura 9 – Escala gradiente de soância dos segmentos. . . . .	34
Figura 10 – Representação da escala de soância nas sílabas de <traço>. N = núcleo da sílaba. . . . .	34
Figura 11 – Representação da escala de soância nas sílabas de <reis>. N = núcleo da sílaba. . . . .	35
Figura 12 – Gráfico comparativo das durações de sílabas pretônicas com a tônica em turco. Fonte: (GORDON; APPLEBAUM, 2010, Fig. 3, p. 43). . . . .	41
Figura 13 – Efeito do acento de palavra na inclinação espectral (medida em <i>phons-dB</i> ) em búlgaro e em macedônio. Fonte: retirado de Crosswhite (2003, p. 769). . . . .	42
Figura 14 – Efeito nulo do acento de palavra na inclinação espectral (medida em <i>phons-dB</i> ) em polonês. Fonte: retirado de Crosswhite (2003, p. 769). . . . .	43
Figura 15 – Diferenças de intensidade de acordo com a tonicidade da vogal. Fonte: retirado de Baleno, Janzen e Yoder (2024, p. 17). . . . .	44
Figura 16 – Diferenças na qualidade vocálica de vogais átonas e tônicas em espanhol. Fonte: retirado de Ronquest (2013). . . . .	45
Figura 17 – Pontos e intervalos intragestuais para um gesto típico. Fonte: retirado de Tilsen (2019, p. 28). . . . .	50
Figura 18 – Efeito de alinhamento de <i>clusters</i> consonantais em torno do núcleo silábico em inglês. (a) = <i>sots vs. spots</i> ; (b) = <i>seek vs. seeks</i> . . . . .	52
Figura 19 – Comparação do efeito de centro-c em onsets e codas. Fonte: retirado de Katz (2010). . . . .	52

Figura 20 – Fluxograma dos processos referentes à metodologia deste trabalho. . . . .	57
Figura 21 – Duração dos áudios gravados de cada sujeito. Estão em rosa os participantes excluídos da análise. . . . .	61
Figura 22 – Trecho de um TextGrid recém-gerado pelo EasyAlign. Legenda: 1 - Camada “ <i>phones</i> ”; 2 - “ <i>syll</i> ”; 3 - “ <i>words</i> ”; 4 - “ <i>phono</i> ”; 5 - “ <i>ortho</i> ”. . . . .	62
Figura 23 – Exemplo de espectrograma de um tepe intervocálico na palavra ‘guaraná’. “4” = [r]. . . . .	66
Figura 24 – Exemplo de espectrograma de um tepe em encontro consonantal na palavra ‘sacra’. “4” = [r]. . . . .	67
Figura 25 – Exemplo de espectrograma de um tepe com vogal de apoio em encontro consonantal na palavra ‘obra’. “4” = [r]. . . . .	68
Figura 26 – Espectrograma do contexto fonológico do logatoma ‘sufras’ com o fenômeno de epêntese vocálica pós-/s/. . . . .	69
Figura 27 – Panorama geral do TextGrid pronto para a extração de dados. . . . .	70
Figura 28 – Ilustração do significado da razão de duração. . . . .	71
Figura 29 – Distribuição da tonicidade dos segmentos. . . . .	74
Figura 30 – Distribuições das durações de segmentos e sílabas no conjunto de dados. . . . .	75
Figura 31 – Distribuição da duração de consoantes e vogais no conjunto de dados. . . . .	76
Figura 32 – Frequências de tipos de segmentos e moldes silábicos no conjunto de dados. . . . .	76
Figura 33 – Resumo dos resultados de alongamento dos segmentos (molde silábico por posição na sílaba). O negrito marca a sílaba tônica. . . . .	78
Figura 34 – Comparação da duração da sílaba em diferentes contextos. . . . .	80
Figura 36 – Comparação da duração de /a/ em diferentes contextos. . . . .	82
Figura 37 – Razão de duração /a/ por molde silábico. . . . .	82
Figura 38 – Efeito do acento de palavra medido em duração bruta e razão de duração, com separação por modo de articulação e vozeamento. . . . .	85
Figura 39 – Gráficos de análise de fricativas. . . . .	85
Figura 40 – Gráficos de análise de oclusivas. . . . .	86
Figura 41 – Gráficos de análise de nasais. . . . .	86
Figura 42 – Comparação entre a duração e a razão de duração do tepe. . . . .	87
Figura 43 – Duração das duas realizações de /s/ de acordo com a tonicidade e a posição na sílaba. . . . .	89
Figura 44 – Comparação entre a duração e a razão de duração do [s] em coda medial. . . . .	90
Figura 45 – Duração de [s] de acordo com a posição da coda na palavra, posição do segmento na sílaba e tonicidade. . . . .	90
Figura 46 – Processo de ressilabificação exemplificado por ‘raspa’ e ‘raspar’. . . . .	94
Figura 47 – Representação da escala de soância nas sílabas de <raspa> após o processo de ressilabificação proposto. N = núcleo da sílaba. . . . .	96

Figura 48 – Análises encontradas ao longo da literatura para explicar encontros-sC iniciais. Da esquerda para a direita: [s] em apêndice à sílaba; [s] em coda de sílaba de núcleo não-realizado; e [sp] como um segmento complexo. Fonte: retirado de Polgárdi (2022, p. 2). . . . .

# Lista de tabelas

Tabela 1 – Sumário de estruturas avaliadas para apontar o mecanismo de manifestação acústica do acento ao longo da literatura. V = vogal; C = consoante; S = sílaba.	23
Tabela 2 – Exemplos de diferentes estruturas silábicas em inglês.	36
Tabela 3 – Alguns tipos de moldes silábicos no português brasileiro e exemplos.	37
Tabela 4 – Correlatos acústicos do acento de palavra segundo diversos autores.	41
Tabela 5 – Tipologia referente ao acento de palavra. Fonte: retirado de Hulst (2014, p. 13).	47
Tabela 6 – Correlatos acústicos do acento de palavras em português brasileiro segundo diversos autores.	48
Tabela 7 – Vogais possíveis em cada contexto tônico do Português Brasileiro.	49
Tabela 8 – Resumo dos processos de sândi vocálico externo no português brasileiro.	49
Tabela 9 – Exemplos de palavras utilizadas no experimento com as oclusivas velares e as fricativas labiodentais. Conjunto de dados completo em Anexo B.	58
Tabela 10 – Configurações do equipamento e software utilizados na coleta de dados.	59
Tabela 11 – Distribuição das ocorrências de segmentos em posições de ataque e coda.	65
Tabela 12 – Descrição dos conjuntos de dados utilizados na análise.	71
Tabela 13 – Descrição dos conjuntos de dados utilizados na análise.	71
Tabela 14 – Medidas de razão ( <i>ratio</i> ) de duração em diferentes estudos. Grifos nossos.	72
Tabela 15 – Palavras adicionadas ao conjunto de dados para avaliação do efeito de coda medial.	73
Tabela 16 – Distribuição das ocorrências de [s] em posições mediais e finais de palavras, considerando a tonicidade.	73
Tabela 17 – Resumo geral dos dados.	74
Tabela 18 – Estatísticas resumidas das variáveis numéricas e categóricas.	74
Tabela 19 – Descrição das Variáveis.	75
Tabela 20 – Resultados do modelo misto para razão de duração das sílabas.	79
Tabela 21 – Resultados dos Contrastes de Tukey (Método de Holm) para o alongamento da sílaba tônica.	79
Tabela 22 – Resultados do modelo misto para razão de duração das vogais.	81
Tabela 23 – Resultados do Teste de Comparações Múltiplas (Tukey) com Ajuste de Holm.	81
Tabela 24 – Resultados do modelo misto para razão de duração das consoantes em primeira posição de ataque.	84
Tabela 25 – Resultados dos Contrastes de Tukey (Método de Holm).	84
Tabela 26 – Resultados do modelo misto para duração do tepe em ramificação de ataque complexo.	87
Tabela 27 – Resultados do modelo misto para razão de duração de [s].	88
Tabela 28 – Resultados do modelo misto para razão de duração de [z].	88

Tabela 29 – Resultados do modelo misto para razão de duração da coda silábica de acordo com a sua posição na palavra. . . . .	89
Tabela 30 – Resultados dos Contrastes de Tukey (Método de Holm). . . . .	89

# Lista de abreviaturas e siglas

A	Ataque ou <i>onset</i> da sílaba
AP	<i>Articulatory Phonology</i>
C	Coda da sílaba
FA	Fonologia Articulatória
N	Núcleo da sílaba
ns	Não-significativo
PB	Português Brasileiro
R	Rima da sílaba
SSP	<i>Sonority Sequencing Principle</i> (SSP)

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>21</b>
<b>1.1</b>	<b>Justificativa e delimitação do problema</b>	<b>22</b>
1.1.1	O foco restrito em vogais de sílabas CV	22
1.1.2	Objetivos	25
1.1.3	Hipóteses do estudo	25
<b>1.2</b>	<b>Breve resumo dos capítulos</b>	<b>28</b>
<b>2</b>	<b>REVISÃO DE LITERATURA</b>	<b>30</b>
<b>2.1</b>	<b>A sílaba</b>	<b>30</b>
2.1.1	Visão geral	30
2.1.2	Teoria de Soância	33
2.1.3	Tipos de moldes silábicos	35
2.1.4	A sílaba em PB	36
<b>2.2</b>	<b>O acento de palavra</b>	<b>38</b>
2.2.1	Problemas na definição	39
2.2.2	O acento nas línguas do mundo	40
2.2.2.1	Manifestação acústica	40
2.2.2.2	Correlatos acústicos de outros tipos de proeminência	46
2.2.2.3	Relevância para a fonologia	46
2.2.3	O acento em PB	47
2.2.3.1	Aspectos acústico-articulatórios	47
2.2.3.2	Processos fonológicos que envolvem o acento	48
<b>2.3</b>	<b>A Fonologia Articulatória</b>	<b>50</b>
2.3.1	Gestos como unidades	50
2.3.2	As dinâmicas de tarefa	50
2.3.3	A sílaba e o efeito Centro-C	51
2.3.4	A proeminência e o modelamento prosódico na FA	53
2.3.5	A Fonologia Acústico-Articulatória: proposta de Albano	55
<b>3</b>	<b>METODOLOGIA</b>	<b>57</b>
<b>3.1</b>	<b>Experimento</b>	<b>58</b>
3.1.1	Materiais	58
3.1.2	Procedimentos de coleta	59
3.1.3	Participantes	60
<b>3.2</b>	<b>Tratamento dos dados</b>	<b>60</b>
3.2.1	Dados desconsiderados	60

3.2.2	Manipulação e etiquetagem dos áudios gravados . . . . .	61
3.2.3	Fenômenos fonológicos não esperados . . . . .	62
3.2.3.1	Casos de vozeamento do [s] em coda . . . . .	63
3.2.3.2	Epêntese segmental . . . . .	65
3.2.4	Processamento de Dados no R . . . . .	68
3.2.4.1	Variável de razão de duração . . . . .	70
3.2.5	Tratamento dos dados de coda medial . . . . .	73
3.2.6	Resumo descritivo . . . . .	74
3.2.7	Avaliação estatística . . . . .	76
<b>4</b>	<b>RESULTADOS . . . . .</b>	<b>78</b>
<b>4.1</b>	<b>Duração da sílaba . . . . .</b>	<b>79</b>
<b>4.2</b>	<b>Duração dos segmentos . . . . .</b>	<b>81</b>
4.2.1	O núcleo da sílaba: /a/ . . . . .	81
4.2.2	As partes não nucleares da sílaba . . . . .	83
4.2.2.1	As consoantes em ataque simples ou ramificado . . . . .	83
4.2.2.2	A consoante em ramificação de ataque: [r] . . . . .	86
4.2.2.3	A consoante em coda: /s/ . . . . .	87
<b>5</b>	<b>DISCUSSÃO . . . . .</b>	<b>91</b>
<b>5.1</b>	<b>A sílaba . . . . .</b>	<b>91</b>
<b>5.2</b>	<b>Os segmentos . . . . .</b>	<b>92</b>
5.2.1	O núcleo da sílaba: /a/ . . . . .	92
5.2.2	As partes não nucleares da sílaba . . . . .	93
5.2.2.1	As consoantes em ataque simples ou ramificado . . . . .	93
5.2.2.2	A consoante em ramificação de ataque: [r] . . . . .	93
5.2.2.3	A consoante em coda: /s/ . . . . .	93
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>99</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>102</b>
	<b>APÊNDICES . . . . .</b>	<b>114</b>
	<b>APÊNDICE A – PARECER CONSUBSTANCIADO - PROJETO APRO-</b>	
	<b>VADO . . . . .</b>	<b>115</b>

<b>ANEXOS</b>	<b>120</b>
<b>ANEXO A – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE) . . . . .</b>	<b>121</b>
<b>ANEXO B – CONJUNTO DE DADOS COMPLETO . . . . .</b>	<b>124</b>
<b>ANEXO C – SCRIPT EM R PARA GERAÇÃO DE APRESENTAÇÕES</b>	<b>130</b>
<b>ANEXO D – SCRIPT EM PRAAT PARA REMOÇÃO DE SEMIVOGAL PALATAL . . . . .</b>	<b>132</b>
<b>ANEXO E – SCRIPT EM PRAAT PARA ORGANIZAR CAMADA DE PALAVRAS DO TEXTGRID . . . . .</b>	<b>133</b>
<b>ANEXO F – SCRIPT EM PRAAT PARA MARCAR O VOZEAMENTO DE [j] . . . . .</b>	<b>134</b>
<b>ANEXO G – SCRIPT EM PRAAT PARA EXTRAIR A DURAÇÃO DE INTERVALOS DOS TEXTGRIDS . . . . .</b>	<b>137</b>

# 1 Introdução

A investigação do acento de palavra em Português Brasileiro (PB) é um campo de estudo que frequentemente converge com questões sobre a sílaba, sua estrutura e seu papel prosódico. A sílaba, definida tradicionalmente como uma unidade formada por uma vogal, possivelmente acompanhada por consoantes adjacentes (e preferencialmente num configuração CV), é considerada um elemento essencial para entender a organização e os padrões fonológicos das línguas. Esse status da sílaba é reforçado não apenas por sua relevância na delimitação de fenômenos como tom e acento, mas também por sua utilidade em descrever restrições fonotáticas (SELKIRK, 1982; KAHN, 1976). Mesmo em modelos menos abstratos, como a Fonologia Articulatória, a sílaba é reconhecida como uma unidade prosódica fundamental, pois organiza segmentos em torno de eventos articulatórios (BROWMAN; GOLDSTEIN, 1986; HALL, 2010).

No que se refere ao acento de palavra, este é entendido como a proeminência atribuída a uma sílaba específica em uma palavra. Tal proeminência não é apenas uma abstração fonológica, mas frequentemente apresenta correlatos acústicos claros, como maior duração, intensidade ou variações de altura tonal (FRY, 1955; LEHISTE, 1970; BECKMAN, 1986). Contudo, as definições de “acento” e sua categorização entre diferentes teorias permanecem desafiadoras, revelando a necessidade de maior rigor terminológico e teórico (HULST, 2014). No contexto do PB, em especial, as sílabas tônicas e átonas desempenham papéis distintos na interação com processos fonológicos, e esses contrastes são centrais para a análise do acento de palavra.

A abordagem adotada nesta pesquisa é fundamentada na Fonologia Articulatória (FA), que considera os gestos articulatórios como unidades representacionais centrais. Esse modelo dinâmico permite capturar variações ao longo do tempo, i.e. de duração, em contextos específicos da sílaba, reconhecendo a influência de fatores de alinhamento gestual dentro da sílaba, e.g. o efeito Centro-C (BYRD, 1995).

O presente trabalho será construído a fim de descrever os efeitos acústicos do acento de palavra do português brasileiro (PB) em **sílabas de diferentes moldes** e em **segmentos de diferentes posições silábicas**. Assim, é explorada a relação entre a estrutura da sílaba, o acento de palavra e os gestos articulatórios no PB, destacando como essas dimensões interagem para moldar padrões duracionais e de alinhamento na língua. Essa abordagem interdisciplinar permite uma análise abrangente, considerando tanto os fundamentos teóricos quanto os dados empíricos, com ênfase na duração como o principal correlato acústico do acento.

## 1.1 Justificativa e delimitação do problema

Duração, intensidade e qualidade vocálica são as principais medidas utilizadas em estudos que buscam compreender o acento de palavra, mas a sua realização acústica exata nas línguas ainda não é totalmente clara para os linguistas. Três de algumas das causas para isso são (i) a falta de transparência quanto aos métodos de coleta e análise de dados por parte de estudos anteriores<sup>1</sup>; (ii) as limitações tecnológicas da época dos estudos-base ao tema (anos 50 em diante) se levarmos em conta os avanços no campo da fonética, fonologia e estatística desde então<sup>2</sup>; e, tratada com mais foco neste trabalho, (iii) a desconsideração do efeito de moldes silábicos complexos e de constituintes não-vocálicos para a realização acústica do acento de palavra.

Por isso, as perguntas que guiam este trabalho são: sílabas com moldes mais complexos são afetadas de maneira diferente pelo acento de palavra? O que acontece com a duração dos segmentos dessas sílabas? As diferentes posições de uma sílaba apresentam os mesmos padrões duracionais/rítmicos?

Nas subseções seguintes, serão detalhados os diferentes moldes silábicos considerados na análise, os segmentos medidos e a justificativa para sua inclusão, bem como as hipóteses do estudo sobre o efeito do acento de palavra na duração dos segmentos silábicos.

### 1.1.1 O foco restrito em vogais de sílabas CV

Estudos que buscaram descrever acusticamente o acento de palavra parecem não ter se ocupado em medir sílabas mais complexas do que a CV (cf. Tabela 1). Embora se espere que a duração de uma sílaba aumente de acordo com seu número de segmentos, ainda não está claro como sílabas de diferentes tamanhos são afetadas pelo acento. Consequentemente, ao ignorar sílabas complexas, a coda, por exemplo, é excluída da análise. Precisamente, nem mesmo o ataque simples é considerado na maioria dos estudos, como será apresentado ao longo desta subseção. Pensamos que a já apontada dificuldade em se atribuir características acústico-fonatórias mais claras ao fenômeno do acento de palavra se deve, ao menos em parte, aos métodos de análise até agora adotados.

Ao focar excessivamente no tipo silábico CV, considerado, por alguns estudos, o único molde existente (LOWENSTAMM, 1996), corre-se o risco de ignorar padrões silábicos mais complexos, como CCV ou CVC, que também desempenham papéis significativos na fonologia de muitas línguas. Essa limitação pode levar a análises incompletas ou enviesadas, subestimando a diversidade de estruturas silábicas existentes. Em 2.1, é apresentada uma discussão sobre o papel da sílaba para os estudos fonológicos.

<sup>1</sup> Por exemplo: ausência de detalhes, dificuldades na replicabilidade, problemas na comparação entre estudos, interpretação ambígua dos resultados.

<sup>2</sup> Por exemplo: equipamentos de gravação de pouca mobilidade e/ou que geravam áudios com baixa qualidade; análise espectral limitada e menos precisa; processamento manual dos dados, que pode levar mais frequentemente a erros; e restrições na quantidade de dados, por dificuldades de seu armazenamento, o que põe em risco a validade da amostra.

Ainda, embora a sílaba tenha se firmado como a unidade de análise pelos estudos fonológicos, em trabalhos cujo foco é a avaliação acústico-articulatória, i.e. fonética, do fenômeno, ainda não há consenso sobre qual(is) unidade(s) devem ser medidas, como se vê na Tabela 1. A rigor, a seleção do(s) segmento(s) avaliado(s) não é justificada. Essas escolhas metodológicas fazem com que “sílaba” e “vogal” sejam termos usados de maneira equivalente: “*Moreover, the average duration of vowels in prominent syllables was approximately 50% longer than the average duration of non-prominent vowels*” (GOEDEMANS; ZANTEN, 2007, p. 42) (grifos nossos); neste outro caso, vemos o descarte da avaliação de consoantes: “*For the study presented here, only the acoustic parameters for vowels were taken into account due to their importance to signal stress in both BP and Swedish*” (BARBOSA; ERIKSSON; ÅKESSON, 2012, p. 4).

Título do Trabalho	Língua(s) analisada(s)	Segmento(s) considerado(s)
Lehiste (1966)	Estoniano	<i>Informação indisponível</i>
Major (1985)	Português Brasileiro	S
Gordon (1995)	Estoniano	C, V
Everett (1998)	Pirahã	C, V
Moraes (1998)	Português Brasileiro	S
Dogil (1999)	Lituano, Polonês, Alemão e Espanhol	C, V
Williams (1999)	Escocês	C
Goedemans e Zanten (2007)	Indonésio	V
Barbosa, Eriksson e Åkesson (2012)	Sueco e Português Brasileiro	V
Barbosa, Eriksson e Åkesson (2013)	Português Brasileiro	V
Vogel, Athanasopoulou e Pinkus (2016)	Espanhol, Grego, Tarco e Húngaro	V
Eriksson et al. (2016)	Italiano	V
Borise (2023)	Georgiano	S
Mousikou, Strycharczuk e Rastle (2024)	<i>Logatomos dissilábicos</i>	S
Severijnen, Bosker e McQueen (2024)	Holandês	S

Tabela 1 – Sumário de estruturas avaliadas para apontar o mecanismo de manifestação acústica do acento ao longo da literatura. V = vogal; C = consoante; S = sílaba.

Verifica-se, conforme ilustrado na Tabela 1, que, dos 16 trabalhos revisados, oito (50%)

medem somente a sílaba como um todo, cinco (aprox. 30%) somente a vogal e apenas quatro (25%) avaliam a consoante individual e acusticamente.

Fry (1955), com o objetivo de analisar os sinais acústicos do acento de palavra no inglês americano, elaborou um experimento para comparar certos pares de nomes e verbos<sup>3</sup>. Especificamente quanto à duração dos segmentos, os resultados mostraram que a duração da vogal diferenciou os dois contextos acentuais em 98% dos casos. Para as consoantes, a duração “[was] not materially affected by the shift of stress” (p. 765), embora não tenham sido apresentadas medidas para exemplificar essa afirmação.

No entanto, é importante citar que já existem casos na literatura que apontam para línguas que fazem a marcação da tonicidade com características acústicas acentuadas de consoantes, tipo de segmento muitas vezes ignorado em avaliações acústicas. Com um estudo em outra língua, Nooteboom (1972) conduziu um experimento semelhante para avaliar o efeito do acento de palavra em consoantes no holandês, utilizando não-palavras. Os resultados indicaram que as consoantes apresentaram um aumento menor na duração em comparação às vogais, mas de forma consistente (HEUVEN, 2018, p. 23). Similarmente ao encontrado por Sluijter e Heuven (1995), a parte da sílaba que mais se alongou foi a rima (núcleo + coda), em oposição ao ataque.

Nesse sentido, há evidências de que o efeito da tonicidade pode se manifestar não apenas na vogal, mas também nas consoantes adjacentes, o que justifica a opção por uma abordagem que considere a sílaba integralmente, como destaca Borise (2023):

*Syllable duration rather than vowel duration alone was measured here because the durational effect of stress may affect the consonant(s) in the stressed syllable, either in addition to or instead of the vowel: e.g., stressed syllables are marked by lengthened onsets in Estonian (Gordon 1995; Lehiste 1966) and codas in Welsh (Williams 1999); for more examples and discussion, see Gordon and Roettger (2017). (BORISE, 2023, p. 15)*

Em estudos preliminares, Nevado (2020), Cantoni e Prado (2023) e Prado e Cantoni (2024) avaliaram a questão no PB. Os seus resultados revelaram que sílabas tônicas tendem a ter consoantes mais longas do que sílabas pós-tônicas. Especificamente, tanto consoantes contínuas (como fricativas e nasais) quanto não-contínuas (como oclusivas) demonstraram um efeito de prolongamento devido ao acento. As duas últimas pesquisas citadas utilizaram dados experimentais coletados de 10 participantes nativos do PB e empregaram modelos mistos lineares para modelar o efeito da tonicidade e estrutura silábica na duração de segmentos, controlando variáveis como a posição na frase e o tipo de consoante. Esses achados sugerem que o papel do ataque não deve ser subestimado – ao menos no PB –, destacando a importância de incluir a medida de consoantes na análise do acento de palavra.

<sup>3</sup> Em inglês, há diversos casos em que sequências de segmentos idênticas têm seus significados diferenciados somente pela posição do acento de palavra, como em *address* (“endereço”, nome) e *address* (“endereçar”, verbo) (o negrito indica a sílaba acentuada). O autor utilizou essa situação para comparar casos em que somente a posição do acento foi modificada, analisando sílabas não acentuadas e seus pares acentuados. As palavras foram inseridas na frase “*where is the accent in...*” (“onde está o acento em...”).

A subestimação do conteúdo consonantal da sílaba tem perpetuado inconsistências na avaliação do acento. A disparidade conceitual e metodológica leva a uma interpretação imprecisa dos correlatos acústicos do acento de palavra, e neste trabalho argumentamos que a delimitação mais precisa dos constituintes não-nucleares, frequentemente marginalizados, pode proporcionar um entendimento mais abrangente dos mecanismos que regem a manifestação do fenômeno.

A não consideração do ataque silábico, por exemplo, carece de embasamento empírico, já que é fruto do entendimento de que esse constituinte silábico não possui a unidade duracional mora ( $\mu$ ) (HAYES, 1989), fato criticado por alguns estudos em diferentes línguas (TOPINTZI, 2010; TOPINTZI, 2011) (mais sobre a unidade  $\mu$  na Subseção 2.1.1). Além disso, pelas condições cinéticas de produção de fala (cada fone produzido tem seu espaço temporal individual, ainda que com algum nível de sobreposição), o aumento do número de segmentos considerados altera as medidas acústicas referentes à duração, que é a característica acústica mais utilizada na maioria das línguas para avaliar a tonicidade de uma determinada sílaba, como será apresentado na Subseção 2.2.2.1.

### 1.1.2 Objetivos

Este estudo busca investigar o papel da duração dos constituintes silábicos (vocálicos e não-vocálicos) na manifestação acústica do acento de palavra no português brasileiro (PB), considerando a interação entre o acento e os diferentes moldes silábicos: CV, CCV, CVC e CCVC.

Os objetivos específicos são:

- examinar, por meio de experimento acústico e análises quantitativas, o papel da duração de cada constituinte silábico (vogal e consoantes) na manifestação do acento de palavra em diferentes moldes silábicos;
- avaliar a possível influência da posição dos constituintes dentro da sílaba (ataque vs. coda) na variação temporal associada ao acento de palavra.

### 1.1.3 Hipóteses do estudo

Este estudo parte da hipótese central de que, no PB, a duração dos segmentos em sílabas acentuadas, tanto vocálicos quanto não-vocálicos, é influenciada por uma interação entre o acento de palavra e o molde silábico. Argumenta-se que sílabas acentuadas compostas por constituintes silábicos simples tenderão a apresentar um alongamento proporcionalmente maior em comparação com sílabas de moldes mais complexos. Este efeito estaria relacionado à maior liberdade temporal de ajustes em regiões específicas da sílaba, assim como aos tipos de segmentos que podem ocupar constituintes específicos.

A hipótese é que, ao se analisar os segmentos silábicos individualmente, será possível observar um aumento significativo na duração dos segmentos não vocálicos (em ataque e coda silábicos) e seu possível impacto na estrutura rítmica da sílaba. Ademais, é esperado que os diferentes moldes silábicos (CV, CCV, CVC e CVCC) tenham impacto no aumento da duração na sílaba acentuada (Figura 1). Essa avaliação é permitida pela metodologia do presente trabalho, que pretende medir todos os segmentos silábicos individualmente, proporcionando um olhar mais detalhado ao papel de cada segmento silábico (Figura 2).

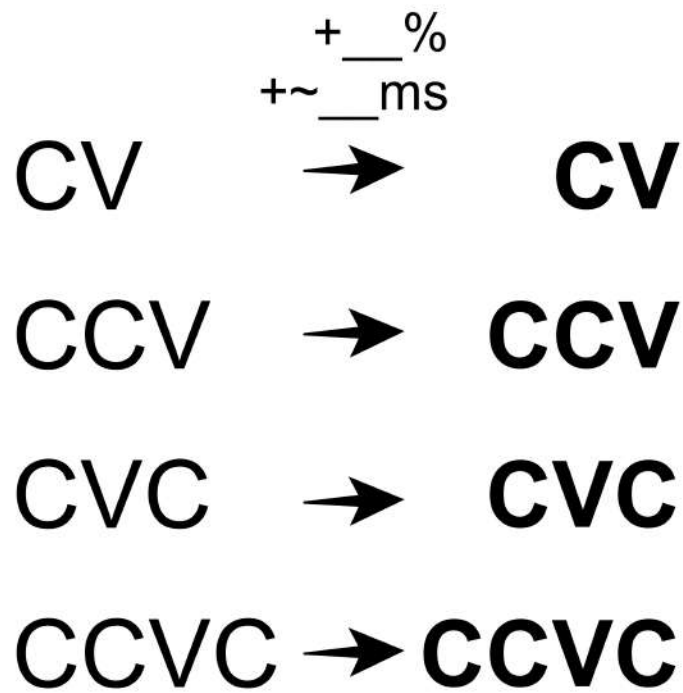


Figura 1 – Proposta de análise de alongamento da sílaba como um todo em quatro moldes silábicos.

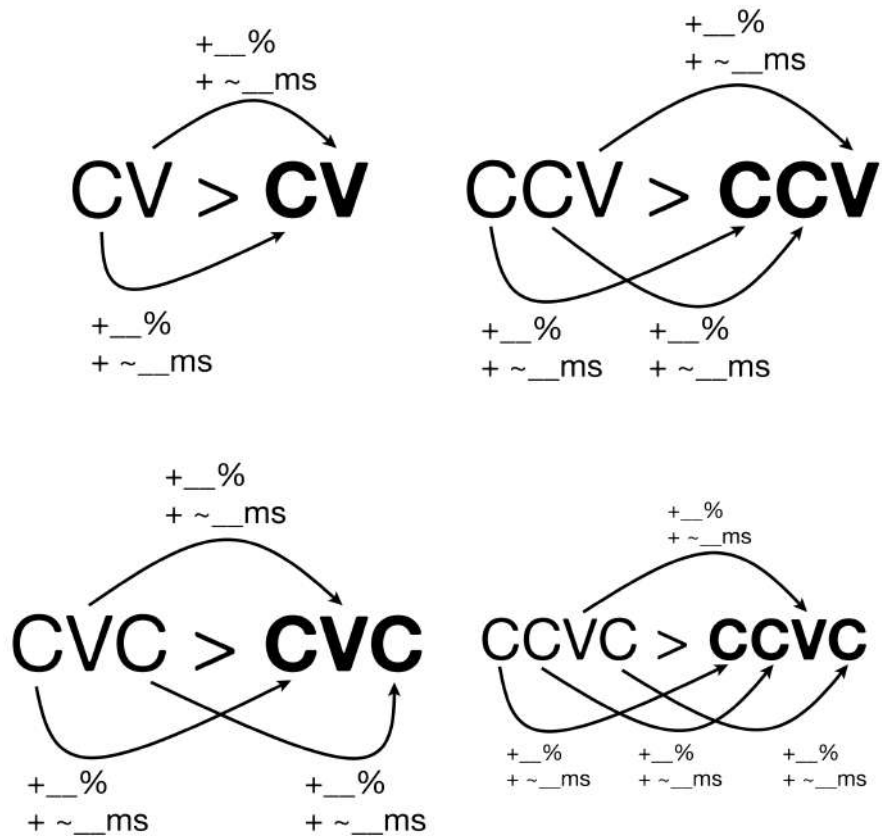


Figura 2 – Proposta de análise de alongamento de cada segmento da sílaba em quatro moldes silábicos.

Em relação às vogais, espera-se que o efeito do acento de palavra promova o seu alongamento em todas as sílabas acentuadas, como já demonstrado em estudos anteriores (ver 2.2.3), e mantendo a mesma proporção de alongamento nos diferentes moldes silábicos. Esse padrão reflete a estabilidade temporal do núcleo silábico, que é a “âncora” rítmica do ataque da sílaba, segundo a Fonologia Articulatória (ver 2.3).

Para as consoantes, sugere-se que aquelas localizadas na primeira posição de ataque de sílabas acentuadas apresentem um efeito de alongamento<sup>4,5</sup>, similar ao observado para as vogais. Este comportamento é especialmente esperado devido ao que é conhecido como efeito Centro-C<sup>6</sup>. Por outro lado, consoantes em coda, por se situarem em uma posição da sílaba menos suscetível a reajustes temporais, devem demonstrar pouco ou nenhum efeito de alongamento, o que reforça a ideia de que a liberdade temporal das codas é menor.

Não é esperado nenhum efeito de acento de palavra ou de molde silábico sobre a duração da consoante que ocupa a ramificação de um ataque complexo, que, no presente trabalho é o [r], também chamado de “tepe”. Esse som corresponde a um movimento rápido, preciso e extremamente curto da língua em direção ao ponto de articulação, funcionando como uma oclusiva muito breve (LADEFOGED; MADDIESON, 1996). Essa característica transitória da articulação do tepe impede que ele seja significativamente prolongado sem comprometer sua identidade fonética. Qualquer tentativa de prolongamento resultaria em outro som, como uma oclusiva alveolar vozeada ([d]) ou uma vibrante alveolar ([r]).

Essas hipóteses, fundamentadas na Fonologia Articulatória, consideram fatores articulatorios, temporais e fonéticos e visam oferecer uma perspectiva mais abrangente sobre a manifestação acústica do acento de palavra no PB. A análise detalhada dos segmentos individuais fará com que seja possível explorar as previsões acima e preencher algumas lacunas de estudos sobre o acento de palavra.

## 1.2 Breve resumo dos capítulos

Esta dissertação está organizada em cinco capítulos, além das referências, apêndices e anexos. O Capítulo 2 traz a revisão da literatura relevante. Inicialmente, são discutidos aspectos fundamentais sobre a sílaba, incluindo sua estrutura, os tipos de moldes silábicos e as especificidades da sílaba no PB. Em seguida, analisa-se o acento de palavra, abordando sua definição, manifestação acústica, e relevância fonológica e características particulares ao PB. Por fim, são

<sup>4</sup> Os resultados desta parte do trabalho já foram apresentados (CANTONI; PRADO, 2023; PRADO; CANTONI, 2024) e embasam a hipótese ao observar o padrão de duração do ataque em um número menor de dados.

<sup>5</sup> Não será traçada nenhuma hipótese quanto à diferença entre as consoantes que ocupam a primeira posição do ataque simples vs. do complexo, devido à falta de base existente tanto na literatura sobre o tema, quanto dentro do modelo em que o presente trabalho se apoia.

<sup>6</sup> O efeito Centro-C (*c-center effect*) descreve como os gestos articulatorios de consoantes no ataque silábico são coordenados temporalmente em relação ao núcleo da sílaba, criando um ponto médio estável, independentemente do número de consoantes no ataque. Esse fenômeno será discutido em maior detalhe na Seção 2.3.3

apresentados conceitos da Fonologia Articulatória, como gestos articulatórios e sua relação com a sílaba e o acento de palavra.

No Capítulo 3, a metodologia adotada na pesquisa é detalhada, incluindo o processo de coleta de dados, os critérios de seleção dos participantes, os procedimentos de gravação e os equipamentos utilizados. Também descreve-se como os dados foram tratados, incluindo etapas de manipulação, segmentação e extração de medidas acústicas, além das estratégias de análise adotadas.

O Capítulo 4 apresenta e analisa os resultados do estudo. Os dados são discutidos à luz das hipóteses propostas, com destaque para os padrões de duração de sílabas e de segmentos.

Por fim, no Capítulo 5, são feitas as considerações finais com as principais contribuições da pesquisa, suas limitações e possíveis desdobramentos para estudos futuros.

## 2 Revisão de Literatura

### 2.1 A sílaba

#### 2.1.1 Visão geral

Como será apresentado na Subseção 2.2.3.1, ao menos em PB, o estudo do acento é mais comumente feito a partir de uma visão fonológica, a qual envolve a consideração de que o escopo de atuação do acento em uma palavra é a sílaba ( $\sigma$ ) (HYMAN, 2006). Por isso nesta seção apresenta-se uma revisão sobre a sílaba para os estudos fonético-fonológicos.

Tradicionalmente, uma sílaba consiste em uma vogal, geralmente precedida por uma ou mais consoantes, e às vezes seguida por uma ou mais consoantes. Embora a literatura que trata da sílaba seja extensa e diferenças teóricas existam, a maioria dos autores concordam, ao menos, que essa unidade existe, principalmente pelas fortes indicações que essa unidade existe, com base em argumentos como:

*First of all, it can be argued that the most general and explanatory statement of phonotactic constraints in a language can be made only via the syllabic structure of an utterance. Second, it can be argued that only via the syllable can one give the proper characterization of the domain of application of a wide range of rules of segmental phonology. And, third, it can be argued that an adequate treatment of suprasegmental phenomena such as stress and tone requires that segments be grouped into units which are the size of the syllable. (SELKIRK, 1982, p. 337).*

Até mesmo autores que costumam evitar abstrações excessivas em suas análises linguísticas concordam que considerar a existência de agrupamentos de segmentos em torno de uma unidade silábica beneficia seus modelos. A Fonologia Articulatória (FA) – ou *Articulatory Phonology* (AP) – (BROWMAN; GOLDSTEIN, 1986) considera a existência dessa unidade prosódica<sup>1</sup>, já que

*[n]early all phonologists agree that syllable structure, or something like it, is an important part of speech planning. Many phonemes have different allophones depending on where they occur in the syllable, and syllables are useful in describing restrictions on how sounds may be combined. (HALL, 2010, p. 822)*

Na maioria das línguas, a sílaba desempenha um papel importante para delimitar o ambiente de atuação de vários fenômenos fonológicos ou de elementos prosódicos (KAHN, 1976), como o tom (KAO, 1971; GUSSENHOVEN, 2004) e o acento de palavra (ver 2.2).

<sup>1</sup> Mais sobre a sílaba no modelo teórico da Fonologia Articulatória na Subseção 2.3.3.

Do ponto de vista descritivo, grande parte da importância da noção de sílaba é o seu uso para esclarecer características da fonotaxe de uma língua, ou seja, os princípios que descrevem quais sequências de sons são encontradas. Ainda, levar em conta a posição do segmento na sílaba torna a descrição da fonotaxe de uma língua mais fácil, já que alguns segmentos são licenciados em final de sílaba, mas não no começo dela, como o [r] em algumas variedades do PB. As restrições fonotáticas estão muitas vezes associadas à estruturação interna da sílaba, cuja concepção varia na literatura sobre fonologia (MENDONÇA, 2003; SCHOOR; AICHERT; ZIEGLER, 2012; DERWING; EDDINGTON, 2014). Ainda que muitas vezes a representação vá de encontro a uma disposição arbórea, comum em modelos gerativos, existem diferenças no que diz respeito a subconstituintes e níveis hierárquicos dentro da sílaba, como se vê a seguir nas Figuras de 3 a 8, onde estão apresentadas algumas das visões correntes na literatura.

1. **Modelo de estrutura linear ou plana:** Neste modelo, a sílaba é vista como uma unidade plana, onde os segmentos que a compõem são organizados linearmente sem ramificações internas (ANDERSON, 1969; KAHN, 1976; CLEMENTS; KEYSER, 1983) (Figura 3). Esse modelo não considera a hierarquia entre os elementos da sílaba e, portanto, não reflete as interações fonológicas e morfológicas que podem ocorrer em línguas com estruturas silábicas mais complexas.

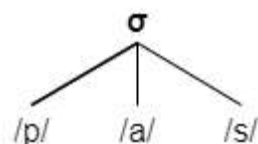


Figura 3 – Representação da estrutura silábica de <paz> de acordo com um modelo de estrutura linear ou plana.

2. **Modelo com ramificação:** Neste modelo, a estrutura silábica é representada com ramificações que conectam o núcleo (vogal) ao ataque (consoantes anteriores) ou à coda (consoantes finais) (PIKE; PIKE, 1947; KURYLOWICZ, 1948; FUDGE, 1969). A ramificação pode ser direcionada para a direita, como mostrado na Figura 4, ou para a esquerda, como na Figura 5. A ramificação reflete uma visão hierárquica da sílaba, onde diferentes partes da sílaba são agrupadas em níveis diferentes. A Fonologia de Governo (Figura 6) sugere que a coda e o ataque compartilham uma relação de governo com o núcleo (KAYE; LOWENSTAMM; VERGNAUD, 1990; HALLE; VERGNAUD, 1987).

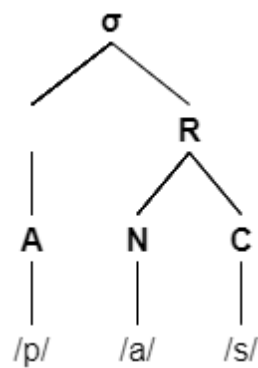


Figura 4 – Representação da estrutura silábica de <paz> de acordo com um modelo de ramificação à direita.

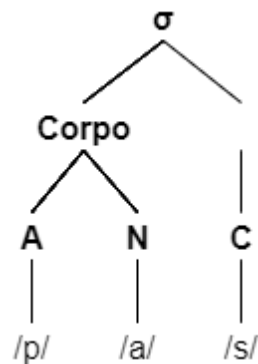


Figura 5 – Representação da estrutura silábica de <paz> de acordo com um modelo de ramificação à esquerda.

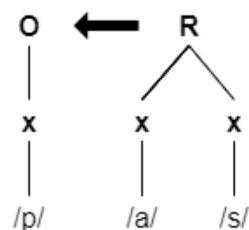


Figura 6 – Representação da estrutura silábica de <paz> de acordo com a Fonologia de Governo.

3. **Modelo com unidade temporal mora:** considera que os segmentos de uma sílaba ocupam *slots* moraicos (HYMAN, 1985; MCCARTHY; PRINCE, 1986; HAYES, 1989). Cada mora representa uma unidade temporal, e.g. uma vogal comum possui uma mora, enquanto as longas são atribuídas duas moras. Esta visão entende que o ataque não é um constituinte

próprio – e, portanto, não atribui peso silábico – e não impõe restrições à complexidade dos *clusters* em posição de ataque (HULST, 1999, p. 30) (mais sobre o tratamento do *onset* na tradição fonológica na Subseção 1.1.1).

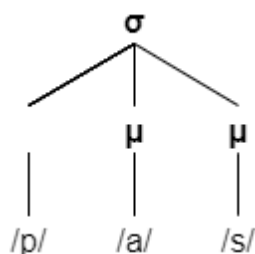


Figura 7 – Representação da estrutura silábica de <paz> de acordo com um modelo moraicista.

4. **Modelo ternário:** Este modelo considera uma estrutura silábica mais complexa, onde a sílaba é dividida em três partes principais: ataque, núcleo e coda, com uma abordagem que pode incluir ramificações internas e múltiplos níveis de representação (HOCKETT, 1955; HAUGEN; HALLE et al., 1956; DAVIS, 1985). O modelo ternário pode acomodar uma maior variedade de padrões silábicos e é útil para descrever línguas que exibem uma rica estrutura interna nas sílabas, permitindo uma análise mais detalhada da interação entre diferentes componentes da sílaba.

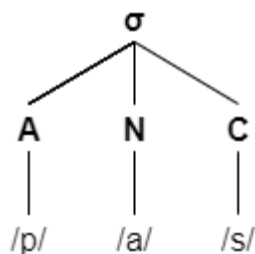


Figura 8 – Representação da estrutura silábica de <paz> de acordo com um modelo ternário.

## 2.1.2 Teoria de Soância

Para além da representação e do entendimento da estrutura interna da sílaba, teorias foram elaboradas para lidar com os tipos de segmentos que ocupam cada porção da unidade. A Teoria de Soância (ver Clements (1990) para uma revisão) é a mais conhecida; Ohala e Kawasaki-Fukumori (1997, p. 319)<sup>2</sup> atribuem a sua primeira aparição ao final do século XVIII, por Brosches (1765, p. 130-133).

<sup>2</sup> Ver Ohala e Kawasaki-Fukumori (1997) para uma crítica à Teoria de Soância, assim como uma alternativa.

A Teoria de Soância<sup>3</sup> propõe níveis hierárquicos de soância. Assim, diferentes tipos de segmentos têm níveis inerentes de soância<sup>4</sup> (Figura 9), influenciando sua compatibilidade com determinadas posições dentro da sílaba. A essa regra, aparentemente comum entre as línguas naturais, é dado o nome de Princípio do Sequenciamento de Soância (*Sonority Sequencing Principle* (SSP)) (SELKIRK, 1984). A soância necessária para ocupar cada um dos constituintes silábicos pode variar ligeiramente entre as línguas, mas é entendido que “*if a language permits syllabic segments from a lower sonority class, it also allows nuclei from all higher sonority classes*” (PARKER, 2011, p. 3). De forma geral, esses contornos contam com picos correspondendo a vogais (núcleos silábicos) e vales indicando limites entre sílabas (ataque e coda) (Figuras 10 e 11). Por fim, pode-se entender que existe uma curva de soância dentro da sílaba como um todo, onde as consoantes em contextos de soância crescente estão no início da sílaba e decrescente em coda da sílaba.

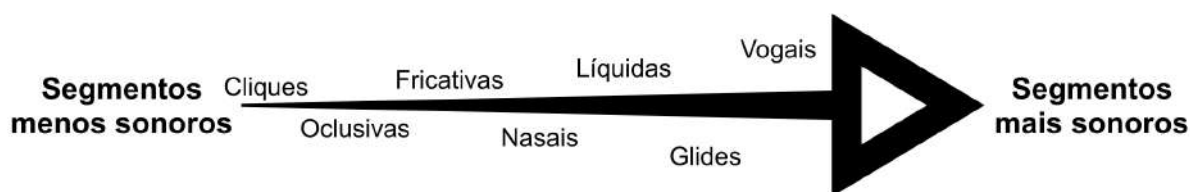


Figura 9 – Escala gradiente de soância dos segmentos.

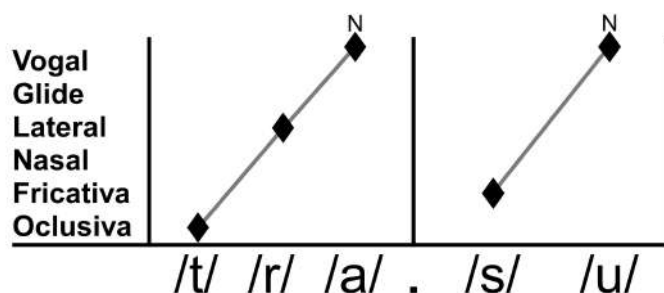


Figura 10 – Representação da escala de soância nas sílabas de <traço>. N = núcleo da sílaba.

<sup>3</sup> Vale ressaltar que há um debate em torno da escala de soância, com propostas alternativas àquela descrita no texto e ilustrada na Figura 9.

<sup>4</sup> Ver uma descrição mais específicas de soância, e.g. vogais baixas são mais sonoras que as altas, em Parker (2008).

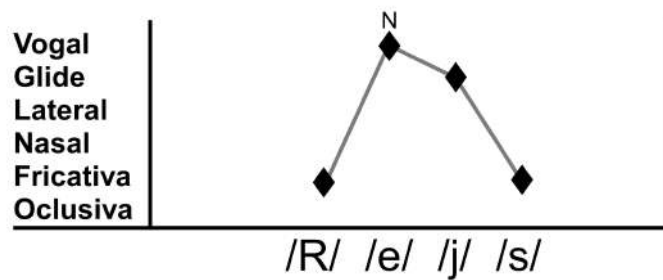


Figura 11 – Representação da escala de soância nas sílabas de <reis>. N = núcleo da sílaba.

### 2.1.3 Tipos de moldes silábicos

As línguas do mundo variam quanto aos tipos de sílabas que aceitam. Algumas, como o Hua (HAIMAN, 1980), possuem apenas um tipo de sílaba: CV. Outras, como o inglês, têm mais de dez formas silábicas básicas (Tabela 2). Os moldes silábicos do PB são apresentados na Subseção 2.1.4. Apesar dessa variação, estudos sobre o tema chegaram a certas generalizações (BLEVINS, 1995):

- Todas as línguas possuem sílabas CV.
- Se uma língua permite *clusters* de  $n$  consoantes no início ou no final de sílabas, ela também permitirá *clusters* de  $n - 1$  consoantes nessas posições.
- Se uma língua não permite sílabas compostas apenas por uma vogal, ela não permitirá sílabas que comecem com uma vogal.

<b>Estrutura Silábica</b>	<b>Exemplos</b>
V	<i>oh, I</i>
CV	<i>to, you</i>
CCV	<i>blue, grow</i>
CCCV	<i>spray, stray</i>
VC	<i>at, up</i>
CVC	<i>cat, back</i>
CCVC	<i>sleep, shrewd</i>
CCCVC	<i>street, splash</i>
VCC	<i>and, eggs, oaks</i>
CVCC	<i>cats, bolt</i>
CCVCC	<i>stoves, flex</i>
CCCVCC	<i>splint, splashed</i>
VCCC	<i>ends, acts</i>
CVCCC	<i>facts, minds</i>
CCVCCC	<i>clasps, sphinx</i>
CCCVCCC	<i>splints</i>
VCCCC	<i>angst</i>
CVCCCC	<i>tempts, sixths</i>
CCVCCCC	<i>branched, twelfths</i>
CCCVCCCC	<i>strengths</i>

Tabela 2 – Exemplos de diferentes estruturas silábicas em inglês.

Embora existam vários moldes possíveis, é consenso que o molde CV é o mais natural e que a vogal é o núcleo silábico esperado (CLEMENTS; KEYSER, 1983; JONG, 2004).

#### 2.1.4 A sílaba em PB

A sílaba máxima em PB pode ser codificada por:

$$\sigma = (C_1C_2)(S_{V_1})V(S_{V_2})(C_3C_4)$$

Onde:

- $C_1, C_2, C_3, C_4$  são consoantes (ataque e coda).
- $V$  é o núcleo vocálico.
- $S_{V_1}, S_{V_2}$  são semivogais/*glides*.

Sendo assim, alguns dos tipos de moldes silábicos existentes no PB são:

O ataque é a parte da sílaba que precede o núcleo e corresponde ao início da sílaba. No PB, o ataque é limitado a no máximo dois elementos consonantais. A combinação dessas consoantes deve seguir regras específicas de soância (cf. Figura 9), que ditam quais grupos são

<b>Tipo</b>	<b>Exemplo</b>
<b>V</b>	<i>é</i>
<b>CV</b>	<i>copo</i>
<b>VC</b>	<i>os</i>
<b>VV</b>	<i>alto</i>
<b>CCV</b>	<i>prato</i>
<b>CVV</b>	<i>pai</i>
<b>CVC</b>	<i>parto</i>
<b>CCVC</b>	<i>entrar</i>
<b>CCVV</b>	<i>grau</i>
<b>CVVC</b>	<i>mais</i>
<b>CVCC</b>	<i>perspectiva</i>
<b>CVVV</b>	<i>Uruguai</i>

Tabela 3 – Alguns tipos de moldes silábicos no português brasileiro e exemplos.

permitidos e quais são excluídos. Os grupos consonantais permitidos são formados com base em características como obstrução e continuidade do som. Por exemplo:

- São permitidas combinações entre uma oclusiva ([v,f, b, p, d, t, g, k]) e uma lateral ([l]).
- Não há ocorrência de alguns encontros consonantais específicos como [dl] e [vl]<sup>5</sup>, ainda que a sequência de soância esteja sendo respeitada.

A coda é a parte da sílaba que vem após o núcleo, correspondendo ao final da sílaba. Ela pode incluir tóticos, a lateral-aproximante [w] ou fricativas, exceto as labiodentais [f, v]. As oclusivas são geralmente proibidas nesta posição. Portanto:

- No final das sílabas, são permitidos os arquifonemas consonantais /S/, /R/, /l/ e /N/, como, respectivamente, em *casca*, *carta*, *calda*<sup>6</sup> e *mantra*<sup>7</sup>.
- Consoantes como [p] e [f] são desfavorecidas na coda e, quando ocorrem, são frequentemente sujeitas a processos fonológicos de inserção de vogal epentética (COLLISCHONN, 2002; CANTONI, 2015), e.g. *internet* > internet[ɪ] e *golf* > golf[ɪ].

O núcleo, em PB, só pode ser preenchido por vogal, e essa configuração é parte fundamental para que uma sílaba seja considerada bem formada. Dois ou mais segmentos vocálicos podem ocorrer: uma vogal e uma semivogal. No PB, os encontros vocálicos são classificáveis em três tipos principais:

### 1. Ditongos:

<sup>5</sup> Como exceção, ocorre em início de palavra no nome 'Vladimir'.

<sup>6</sup> Situação quando se realiza o fonema como a lateral [l]

<sup>7</sup> Ver discussão sobre a consoante e as vogais nasais no PB em (MEDEIROS, 2007)

- **Ditongo Crescente:** Formado por uma vogal semifechada (ou semivogal) seguida por uma vogal aberta. Exemplo: *pai* ['paɪ], *quase* ['kwa.zɪ].
- **Ditongo Decrescente:** Formado por uma vogal aberta seguida por uma semivogal. Exemplo: *rei* ['fɛj], *pão* ['põw]<sup>8</sup>.

## 2. Hiatos:

- Ocorrência de duas vogais em sílabas separadas. Exemplo: *raís* [fia.'is], *caído* [ca.'i.du].

## 3. Sequências de mais de duas vogais:

- Sequências de três vogais na mesma sílaba são mais raras, geralmente compostas por uma semivogal, uma vogal e outra semivogal. Exemplo: *averigui* [a.ve.ɾi.'gwej], *Uruguai* [u.ru.'gwaj].

É digno de nota que alguns fenômenos e reajustes fonético-fonológicos, como apagamento vocálico ou ressilabificação, podem fazer com que o falante pronuncie sílabas que não seriam esperadas de acordo com a fonotaxe do PB, e.g. sílabas sem núcleo (sem vogal) ou com sequências de três ou mais consoantes (CÂMARA JÚNIOR, 1970; BISOL, 2010a).

É importante destacar que o PB, devido à sua complexidade silábica e diversidade de padrões fonológicos, oferece uma base rica para o estudo de fenômenos acústicos e prosódicos em diferentes tipos de moldes silábicos. Essa plasticidade na estrutura silábica torna o PB uma língua de particular interesse para investigações sobre o papel da sílaba como unidade prosódica e para a análise, por exemplo, de suas interações com o acento lexical. O presente trabalho tem interesse nas interações entre estrutura fonológica e realização acústica do acento de palavra no PB.

## 2.2 O acento de palavra

O acento de palavra (*word stress*) representa a ênfase ou proeminência relativa atribuída a uma sílaba específica em uma palavra ou a uma palavra particular em uma frase (GORDON; HULST, 2020; FRY, 1955; FRY, 1958; LEHISTE, 1970; GOFFMAN; MALIN, 1999; CUTLER, 2005; CUTLER; JESSE, 2021) e é obrigatório e culminativo (HYMAN, 2006, p. 231), ou seja, existe apenas um único em cada palavra, ou em cada *palavra fonológica*. Uma definição exata do termo é uma tarefa em andamento, como avaliado por Hulst (2014), que mostra as diferentes definições adotadas ao longo dos anos por diversos autores (ver Subseção 2.2.1, a seguir).

<sup>8</sup> A natureza fonológica da vogal nasal no PB possui uma extensa literatura com uma análise própria e com diversas visões sobre sua natureza ([ẽ] vs. [aN]) (ABAURRE; PAGOTTO, 1996; CÂMARA JÚNIOR, 1982; SOUSA, 1994). Como esse tipo de segmento vocálico não faz parte do objeto de estudo (cf. Seção 3), não nos aprofundaremos na discussão. Ver Medeiros (2007) para uma revisão.

O presente trabalho tratará de “acento” como uma proeminência acústico-articulatória dada a uma sílaba, que faz com que ela passe por um processo de alongamento<sup>9</sup> (dentre alguns outros, que são menos relevantes para a presente pesquisa). O acento atribuído à sílaba no nível da palavra é denominado acento lexical ou acento de palavra e é o foco do presente trabalho. Para além, existe o acento relacionado a uma estrutura prosódica maior: “*stress is accentuation of syllables within words, or of words within sentences*” (CUTLER, 2005, p. 264). O acento atribuído ao nível da frase é conhecido como acento de sentença e engloba o acento frasal (que destaca a sílaba mais forte na frase) e o acento contrastivo (utilizado para enfatizar um item, palavra ou morfema), e.g.: “**Eu** não quebrei o vaso.” (“Outra pessoa quebrou.”); “Eu não **quebrei** o vaso.” (“Eu fiz outra coisa com isso.”); “Eu não quebrei o **vaso**.” (“Eu quebrei outra coisa.”). A proeminência ao nível da palavra e aquela ao nível da frase possuem, geralmente, diferentes correlatos acústicos, como será apresentado a seguir na Subseção 2.2.2.3.

Em oposição a uma sílaba acentuada, existem as sílabas que não recebem o acento, denominadas átonas. Elas ocupam posições pré-tônicas (antes do acento da palavra) ou pós-tônicas (após), onde regras fonológicas específicas podem se aplicar (ver Subseção 2.2.3.2).

## 2.2.1 Problemas na definição

Embora o conceito de acento seja amplamente discutido na literatura, sua definição precisa ainda permanece um desafio. Hulst (2014) explora essa questão e ressalta como diferentes abordagens variam ao longo do tempo, apontando que

*[...] the use of the terms ‘stress’ and ‘accent’ is somewhat problematic. This may easily lead to confusion when comparing different traditions or theories. In one respect, the two terms can be understood as being translations of each other (as in stress being an English term and accent a French term for the same thing, whatever that thing is). (HULST, 2014, p. 4)*

O termo “acento de palavra” não é consistentemente definido ao longo da literatura, refletindo variações consideráveis tanto entre línguas quanto entre teorias linguísticas. Enquanto alguns autores tratam o acento como uma marcação lexical abstrata que indica a posição de uma sílaba proeminente, outros o associam diretamente a características fonéticas, como duração, intensidade e altura tonal. Abercrombie (1976) sugere que o acento não possui conteúdo fonético intrínseco, sendo uma noção puramente lexical, enquanto outros, como Beckman (1986), enfatizam a distinção entre acento de *pitch* e acento de intensidade. Essa falta de consenso cria dificuldades na comparação de descrições prosódicas entre línguas e contribui para

<sup>9</sup> Questiona-se na literatura questionando se é mesmo a sílaba tônica que se alonga, ou se são as outras sílabas não-tônicas que sofrem um encurtamento. Essa reflexão não será abordada no presente trabalho pela falta de espaço para uma questão complexa e pouco relevante para este trabalho. A escolha por tratar como se o caso fosse de um alongamento da sílaba tônica segue uma tendência geral da literatura sobre acento e tem como finalidade simplificar a discussão que será posta ao longo deste trabalho, além de tornar os resultados mais facilmente comunicáveis.

a ambiguidade na categorização de fenômenos como o *stress-accent* e o *pitch-accent*. Assim, existe a necessidade de maior rigor terminológico e teórico na análise do acento de palavra.

## 2.2.2 O acento nas línguas do mundo

### 2.2.2.1 Manifestação acústica

Após os estudos pioneiros de Fry (1955), Fry (1958), que avaliaram acusticamente o acento de palavra em inglês<sup>10</sup>, outros trabalhos em diferentes línguas surgiram. Os resultados desde então apontam para que a realização dessa característica prosódica seja específica para cada língua, podendo envolver (de maneira individual ou acumulada): inclinação espectral, intensidade, duração, frequência fundamental, entre outros mecanismos. Tais mecanismos para a manifestação acústica fazem com que, do ponto de vista da percepção, as sílabas acentuadas sejam frequentemente percebidas como mais salientes do que as não acentuadas (CUTLER, 2005; CUTLER; JESSE, 2021). O acento é, ainda, conceituado por alguns estudiosos como uma propriedade estrutural que permite que sílabas carreguem um acento tonal, e não necessariamente uma característica puramente fonética (BOLINGER, 1965; BECKMAN; EDWARDS, 1994; TILSEN, 2019). Por se tratar de um fenômeno tão complexo e multifacetado, “[i]t has [...] become increasingly clear that the phonetic study of stress is a complicated matter for several reasons” (GORDON, 2011), mas, de maneira geral,

*[s]tress or lack of stress is associated with both suprasegmental and segmental properties. On a suprasegmental level, stress typically triggers lengthening, higher fundamental frequency, and greater intensity, although there are many languages in which these properties do not converge on a single syllable but rather are distributed over multiple syllables. On a segmental level, stress characteristically, although not always, triggers consonant fortition or blocks lenition targeting unstressed syllables. (GORDON, 2011, p. 20)*

Burroni e Tilsen (2022), em um estudo sobre choque acentual (*stress clash*<sup>11</sup>), tratam dos correlatos acústicos do acento de palavra, e apontam algumas questões e problemas, como: (1) a necessidade de hipóteses que liguem entidades fonológicas, como acento, acentos tonais e fronteiras prosódicas, a medições fonéticas; e (2) a dificuldade de interpretar “correlatos fonéticos” como relações causais, uma vez que essas associações são generalizações estatísticas e, portanto, imprecisas. Cada tipo de correlato acústico será explorado em mais detalhes a seguir<sup>12</sup>.

<sup>10</sup> O trabalho de Okobi (2006), que avalia a manifestação do acento de palavra em inglês, também é digno de nota.

<sup>11</sup> Ver discussões e análises sobre o choque acentual nas línguas em Burroni e Tilsen (2022).

<sup>12</sup> Uma medida não discutida a seguir, por serem poucos os trabalhos que a avaliam, foi o tempo de início de vozeamento (VOT). De maneira resumida, alguns estudos apontam que, para algumas línguas, o VOT é mais longo em sílabas acentuadas (KEATING, 1984; CHO; MCQUEEN, 2005; COLE et al., 2007; LISKER; ABRAMSON, 1967).

Autor	Incl. Espec.	Intens.	Duração	F0	Qualidade Vocálica
Fry (1955)		✓	✓	✓	
Fry (1958)		✓	✓	✓	
Gordon e Roettger (2017)			✓		
Cutler (2005)			✓		
Cutler e Jesse (2021)			✓		
Hulst (2014)	✓	✓	✓	✓	✓

Tabela 4 – Correlatos acústicos do acento de palavra segundo diversos autores.

## Duração

Em um levantamento com cerca de 70 línguas, [Gordon e Roettger \(2017\)](#) apontam que a duração é a principal marca acústica para diferenciar sílabas acentuadas de não-acentuadas em 90% dos casos analisados. Por esta razão, essa marca acústica é aquela mais escolhida para apontar características de produção de sílabas acentuadas. Ver exemplo da duração de uma sílaba acentuada em comparação às outras na [Figura 12](#).

Ainda neste estudo, os autores verificam que não são somente as vogais que são alongadas (p. 5). Os achados variam entre as línguas: em estoniano e espanhol peninsular, as consoantes de maneira geral em sílabas tônicas foram alongadas, enquanto no galês somente as codas. Em Lakhota, embora as vogais não tenham sido alongadas sob ação do acento, o tempo de início de vozeamento (*voicing onset time* (VOT)) para oclusivas aspiradas foi maior em sílabas tônicas; os mesmo segmentos em Yakima Sahaptin tiveram VOTs mais curtos. Alguns estudos também indicaram um efeito geral de alongamento na rima da sílaba acentuada ou da sílaba inteira.

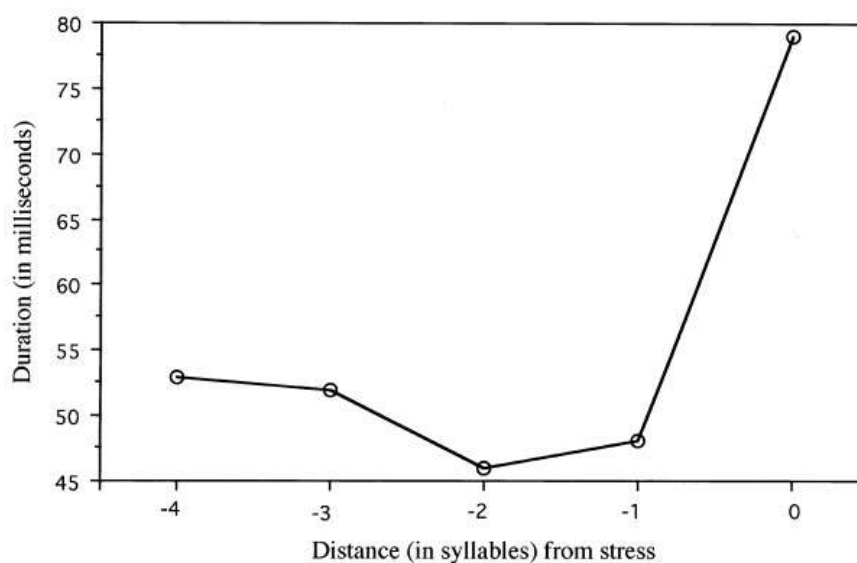


Figura 12 – Gráfico comparativo das durações de sílabas pretônicas com a tônica em turco. Fonte: ([GORDON; APPLEBAUM, 2010](#), Fig. 3, p. 43).

## Inclinação espectral

Alguns estudos indicam que a inclinação espectral (em inglês, *spectral tilt*) pode ser um marcador confiável de acento. Conforme os resultados de [Sluijter e Heuven \(1996\)](#), sílabas acentuadas apresentam maior intensidade nas bandas de frequência acima de 500 Hz, enquanto as sílabas não acentuadas apresentam uma menor intensidade das frequências mais altas. Resultados semelhantes foram encontrados por [Crosswhite \(2003\)](#) no búlgaro e no macedônio (Figura 13), porém não em polonês (Figura 14). Essa diferença de intensidade nas regiões superiores do espectro pode contribuir, segundo os autores, para a percepção de maior proeminência, atribuindo às frequências mais altas a manifestação acústica do acento de palavra.

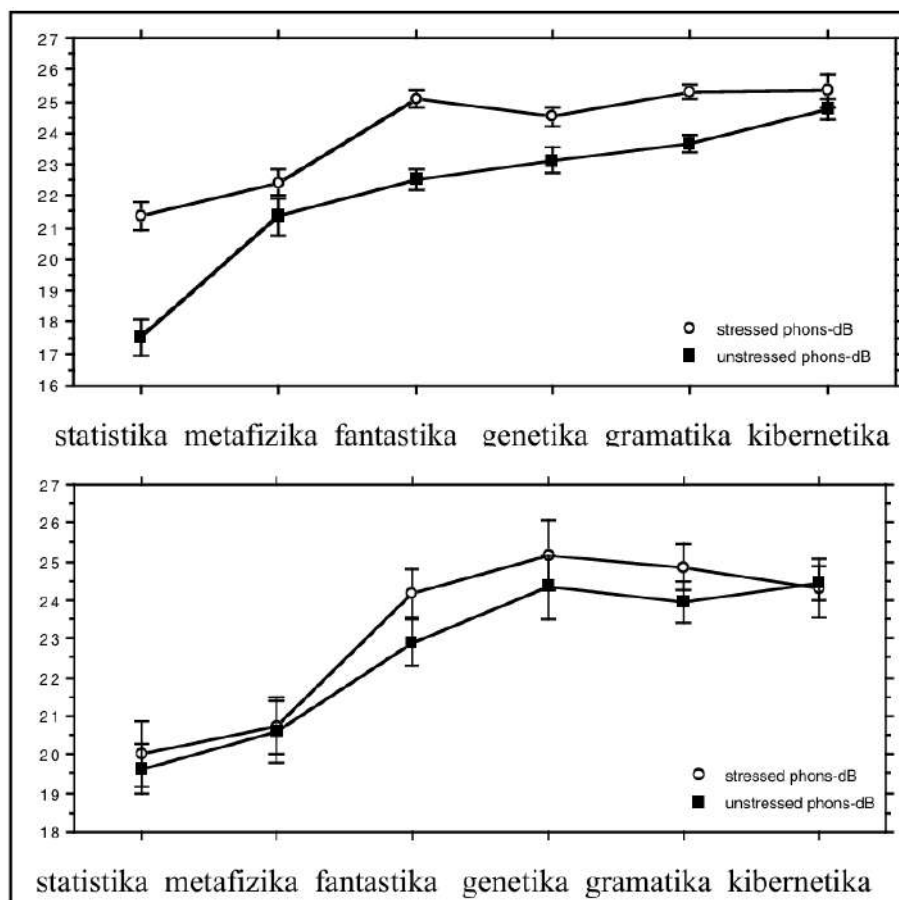


Figura 13 – Efeito do acento de palavra na inclinação espectral (medida em *phons-dB*) em búlgaro e em macedônio. Fonte: retirado de [Crosswhite \(2003, p. 769\)](#).

No entanto, mais estudos são necessários para entender como a inclinação espectral contribui para a realização do acento de palavra. Em algumas línguas, o acento está associado a uma inclinação espectral maior, ou seja, uma queda menos acentuada da intensidade em frequências mais altas. Porém,

[...] no single way of measuring spectral tilt has been definitively shown to correlate with stress vs. vowel quality, or even with stress vs. phrasal-accents (Gordon & Roettger, 2017). ([BURRONI; TILSEN, 2022](#))

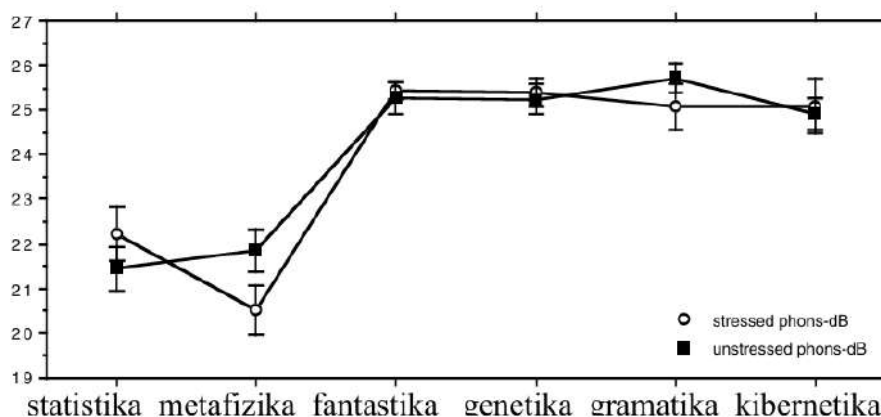


Figura 14 – Efeito nulo do acento de palavra na inclinação espectral (medida em *phons-dB*) em polonês. Fonte: retirado de Crosswhite (2003, p. 769).

### Frequência fundamental (F0)

Segundo Gordon e Roettger (2017), medir a frequência fundamental (F0) de sílabas acentuadas é complexo devido à sobreposição com eventos tonais de nível mais alto prosodicamente, o que exige maior cautela na elaboração dos desenhos experimentais de coleta<sup>13</sup> e na interpretação dos dados. A F0 marcou o acento em 73% das línguas-alvo de seu conjunto de dados, com uma taxa de sucesso maior ao excluir línguas tonais. A maioria dos estudos utilizou medidas estáticas, como a F0 média, enquanto algumas línguas exigiram medidas de contornos dinâmicos para detectar o acento. Exceções foram observadas no basco de Goizueta, onde a F0 distingue o acento apenas em palavras sem tom lexicalmente especificado, e no náhuatl de Balsas, onde a F0 continua a ser indicativa do acento, apesar de ser uma língua tonal. Em geral, as sílabas acentuadas mostraram valores de F0 mais altos, embora alguns estudos tenham relatado uma F0 reduzida indicando acento.

A fim de auxiliar na separação dos dois fenômenos prosódicos, cabe citar que os correlatos acústicos do acento no nível da frase diferem dos correlatos do acento no nível da palavra. Alguns autores consideram a F0 como manifestação acústica de um acento frasal, e não daquele que é foco do presente trabalho. Uma breve discussão sobre isso é apresentada na Subseção 2.2.2.2. Isso fica claro em estudos que comparam sílabas acentuadas em palavras com e sem foco/ênfase contrastiva:

*[...] [I]nvestigation of stressed syllables in and outside of focus has shown that these pitch properties are in fact very often not part of the set of word-level stress properties. When stressed syllables are measured in out-of-focus position they often do not include pitch as a significant factor, but rather comprise primarily the various consequences of articulatory force or hyperarticulation which typically enhance intensity ('loudness'), duration, fullness of articulation*

<sup>13</sup> Vogel, Athanasopoulou e Pinkus (2016) fazem uma discussão sobre a importância de evitar o foco contrastivo em experimentos voltados para a análise do acento de palavra. Ver, ao final da Subseção 2.2.2.1, uma discussão de Borise (2023) sobre a interação entre proeminências de diferentes níveis em estudos sobre o acento de palavra no georgiano.

(with consequences for vowel quality and phonation), and more technical notions such as spectral tilt (or spectral balance), not excluding somewhat elevated pitch, but not the kinds of pitch movements which are introduced by the intonational system as markers of focus (and domain edges) (see Beckman 1986 and Gordon 2011 for relevant discussion and references). (HULST, 2014, p. 7)

## Intensidade

Gordon (2011) indica que a intensidade pode ser considerada um marcador acústico confiável para a identificação do acento. Segundo o autor, a intensidade pode interagir com outros fatores, como a frequência fundamental e a duração, na manifestação do acento de palavra. Beckman (1986), por exemplo, apresenta dados que indicam que, para a veiculação de uma sílaba proeminente na palavra, a intensidade é um fator significativamente relevante em inglês “but not in Japanese, a language employing tone-based lexical contrasts” (GORDON, 2011, p. 3). A Figura 15 exemplifica como seria a intensidade em uma sílaba acentuada.

Mol e Uhlenbeck (1955) discutem a relevância da intensidade na veiculação do acento de palavra, apresentando uma revisão crítica de como a intensidade se relaciona com o acento em comparação com outros indicadores fonéticos. Os autores indicam que Franz Saran estava correto (p. 14) ao argumentar que a intensidade não deve ser considerada um fator primário na análise do acento, porque o ouvido humano é menos sensível a variações de volume do que a diferenças de duração (SARAN, 1907). Por outro lado, outros estudiosos criticam essa visão, defendendo que a intensidade ainda impacta a percepção do acento, e que fatores dinâmicos e musicais contribuem para a sua manifestação acústica (MEYER, 1920; SCHMITT, 1924).

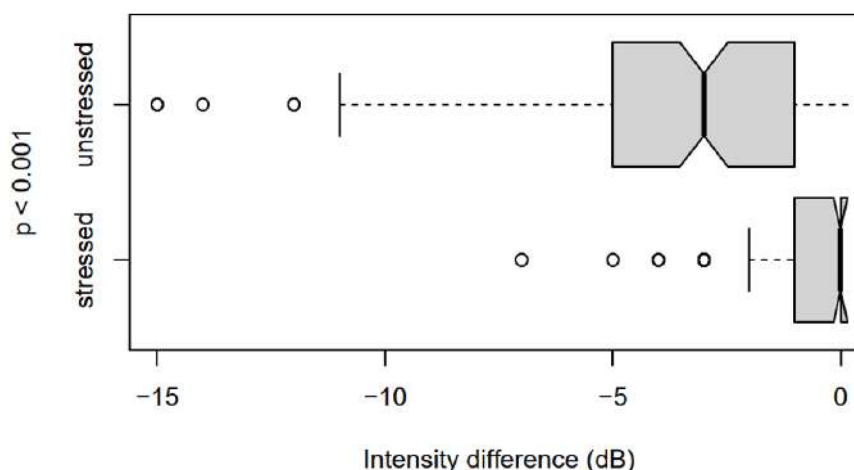


Figura 15 – Diferenças de intensidade de acordo com a tonicidade da vogal. Fonte: retirado de Baleno, Janzen e Yoder (2024, p. 17).

## Qualidade vocálica

A qualidade da vogal acentuada/não-acentuada também pode servir para apontar a existência de uma proeminência acentual, especialmente em línguas que apresentam redução vocálica em posições átonas, como o PB (mais sobre os efeitos do acento de palavra no inventário vocálico do PB na Subseção 2.2.3.2). Segmentos vocálicos acentuados tendem a ser realizados com uma “*fullness of articulation*” (completude de articulação) (HULST, 2014, p. 7). Esse comportamento faz com que, ao se comparar equiláteros com informações de F1 e F2 de vogais, as tônicas ocupem posições mais periféricas no gráfico, e as átonas se mantêm no centro, conforme a Figura 16.

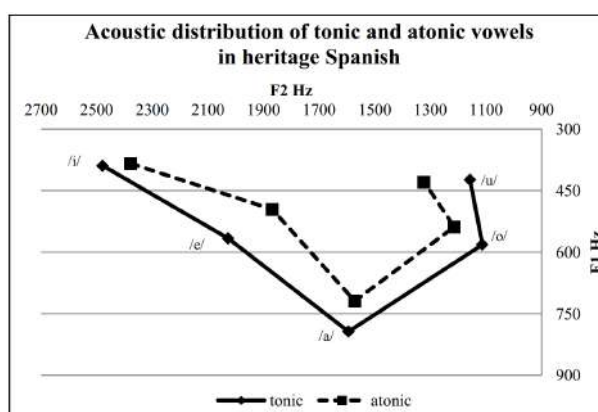


Figura 16 – Diferenças na qualidade vocálica de vogais átonas e tônicas em espanhol. Fonte: retirado de Ronquest (2013).

Esse cenário se refere à frequente redução vocálica que sofrem as vogais pretônicas e postônicas. O fenômeno se refere à tendência de certos sons vocálicos sofrerem mudanças qualitativas em sílabas não acentuadas, muitas vezes resultando na neutralização de contrastes fonéticos que seriam perceptíveis em sílabas tônicas (CROSSWHITE, 2004). É comum, portanto, que vogais em sílabas não-acentuadas caminhem em direção ao [ə], porém

*the extent of vowel reduction in unstressed or unaccented syllables varies among languages. Unstressed vowels are more centralized in English than in Dutch, although they exhibit similar durational patterns (van Bergem, 1993; Sluijter & van Heuven, 1996) whereas in other languages with lexical stress like Arabic, unstressed vowels are not as reduced or short as unstressed vowels in English (e.g., de Jong & Zawaydeh, 2002; Zuraiq & Sereno, 2007). [...] Polish reportedly does not reduce vowels in unstressed position (Jassem, 1959), although some studies have found evidence of vowel formant undershoot and spectral tilt differences in unstressed syllables (e.g., Crosswhite, 2003a; Nowak, 2006). (FLETCHER, 2010, p. 534)*

A qualidade vocálica passa, então, a se tornar também o foco de estudos dedicados à percepção da fala. De acordo com a teoria da dispersão, formulada por Lindblom (1986), as vogais “de canto” (*corner vowels*) do triângulo vocálico — /i/, /u/ e /a/ — são as mais comuns e perceptivelmente distintas devido à sua posição nos extremos do espaço vocálico. Essas vogais

tendem a resistir mais à redução, uma vez que são mais fáceis de serem percebidas em qualquer posição prosódica. Por outro lado, vogais médias, como /e/ e /o/, são mais suscetíveis à redução em sílabas átonas, aproximando-se dos sons mais centrais e neutros. Esse padrão de redução vocálica pode ser visto como uma estratégia de contraste acústico, onde as vogais tônicas mantêm sua plena realização para garantir a clareza perceptiva nas posições mais importantes prosodicamente.

### 2.2.2.2 Correlatos acústicos de outros tipos de proeminência

Por fim, cabe citar que os correlatos acústicos do acento no nível da frase geralmente envolvem a F0, em contraste com o nível da palavra, para o qual a duração é predominante. As curvas de F0, ou a variação de altura tonal, são usadas para marcar o foco e os limites de domínio na frase e é veiculado acusticamente pela frequência fundamental (F0).

É relevante citar como medir a manifestação acústica do acento de palavra pode ser difícil se considerarmos que outras ferramentas linguísticas podem utilizar os mesmos meios para sua materialização no sinal sonoro. Borise (2023) aborda o problema da interação entre o acento de palavra e o frasal em línguas como o georgiano, onde a delimitação clara entre esses dois fenômenos pode ser complexa. No caso do georgiano, existe um debate sobre a existência e a natureza do acento de palavra, com diferentes estudos apontando para diferentes sílabas como portadoras de acento (primeira sílaba (TSCHENKELI, 1958; TEVDORADZE, 1978), antepenúltima (AKHVLEDIANI, 1949; GUDAVA, 1969) ou penúltima (ZHGHEINTI, 1958)). Além disso, a correspondência frequente entre palavras lexicais e frases acentuais (*Accental Phrases* (AP)) dificulta a distinção entre características prosódicas ao nível da palavra e da frase. Os seus resultados mostram que o acento de palavra no georgiano é fixo na primeira sílaba da palavra, porém com manifestação acústica altamente variável, dependendo da posição silábica e da estrutura prosódica das frases.

Estudos que se propõem a examinar o acento de palavra ou o frasal precisam, portanto, considerar as interações entre mecanismos linguísticos de diversos níveis, difíceis de se separar, pois eles podem compartilhar formas de materialização acústica.

### 2.2.2.3 Relevância para a fonologia

Além de características suprasegmentais, a tonicidade de uma sílaba pode desencadear processos fonético-fonológicos em vogais e consoantes. Por exemplo, no inglês americano, os segmentos [t] e [d] sofrem lenição em posição pós-tônica, transformando-se em um tepe alveolar ([ɾ]). Em muitas línguas, como o russo (BARNES, 2007), o português (WETZELS, 1992; WETZELS, 2011) e o inglês (FOURAKIS, 1991), a redução de vogais pode ocorrer quando uma vogal muda de uma posição tônica para uma posição átona. No inglês, vogais não acentuadas podem reduzir-se a vogais semelhantes ao *schwa* ([ə]). De maneira geral, sílabas acentuadas passam por processos de fortificação e de alongamento, enquanto as não acentuadas

sofrem lenição e encurtamento (GORDON, 2011). Assim, podemos vincular segmentos em posições específicas à manifestação acústica do acento em diversas línguas.

Além de avaliar e atestar os usos citados, alguns trabalhos de análise fonológica buscam tratar da (im)previsibilidade de ocorrência do acento. Quando previsível, o acento pode seguir padrões específicos para cada língua, como na primeira sílaba da palavra em finlandês, na segunda em polonês, na última em turco ou, no caso do PB, em uma das três últimas sílabas da palavra. A depender da língua, o peso silábico pode ser um atrator para o posicionamento do acento em uma palavra (GORDON, 2006; GOEDEMANS; HULST, 2013). Normalmente, a fixação do acento de palavra em uma sílaba e a sua sensibilidade ao peso silábico são duas condições utilizadas para a tipologia linguística (HULST, 2014) (Quadro 5).

	<i>Weight-sensitive</i>	<i>Weight-insensitive</i>
<i>Bounded</i>	English	Finnish
<i>Unbounded</i>	Amele	Turkish

Tabela 5 – Tipologia referente ao acento de palavra. Fonte: retirado de Hulst (2014, p. 13).

Algumas línguas, coerentemente aquelas que não possuem acento de palavra fixo, usam a acentuação de uma sílaba para distinguir sequências de segmentos que, sem tal variação, seriam idênticas, como em PB: ['sa.bjɐ] (<sábia>), [sa.'bi.ɐ] (<sabia>) e [sa.bi.'a] (<sabiá>); e em diversas outras línguas românicas, como o espanhol (EDDINGTON, 2004) e o italiano (D'IMPERIO; ROSENTHALL, 1999).

Paralelamente, Laver (1994) indica uma tendência de que línguas tenham seus acentos nas fronteiras de suas palavras, o que aponta para uma função de marcação de fronteiras prosódicas. Ver Thiessen e Saffran (2003) para uma análise com o inglês; e CÂMARA JÚNIOR (1975) para uma com o português.

## 2.2.3 O acento em PB

### 2.2.3.1 Aspectos acústico-articulatórios

Cada língua pode manifestar acusticamente o acento de palavra à sua maneira: aumento de duração (português brasileiro e estoniano), variações de F0 (italiano), intensidade (húngaro e papiamento), inclinação espectral (espanhol peninsular e pitjantjatjara) etc. (GORDON; ROETTGER, 2017). Quanto ao estudo das características de produção do acento específicas do PB, é importante entender que, embora seja um tema há muito estudado (MORAES, 1986; MASSINI-CAGLIARI, 1992; MAJOR, 1981), a maioria dos trabalhos

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

<b>Autor</b>	<b>Duração</b>	<b>F0</b>	<b>Intensidade</b>	<b>Qualidade Vocálica</b>
Fernandes (1977)	✓	✓	✓	-
Major (1985)	✓	-	-	-
Moraes (1998)	✓	✓	✓	-
Ferreira Neto (2007)	✓	✓	✓	✓
Migliorini e Massini-Cagliari (2010)	✓	-	-	-
Cantoni (2013)	✓	-	✓	✓

Tabela 6 – Correlatos acústicos do acento de palavras em português brasileiro segundo diversos autores.

no PB, são majoritariamente devotados à entonação e ao ritmo. (CANTONI, 2013, p. 55)

Dentre os estudos que se debruçaram sobre o aspecto acústico do acento de palavra em PB, a grande maioria aponta que a duração seja o principal recurso utilizado pelos falantes para marcar o acento de palavra (MAJOR, 1985; MIGLIORINI; MASSINI-CAGLIARI, 2010), em consonância com a já citada (2.2.2.3) tendência natural apontada por Gordon e Roettger (2017). O aumento na intensidade sonora e a hiperarticulação segmental (*fullness of articulation*) desempenham também um certo papel na pronúncia do acento de palavra, embora a nível menor, se comparado ao da duração (CANTONI, 2013). Variações na frequência fundamental (F0) estão mais associadas ao acento frasal, e não ao do nível da palavra (MASSINI-CAGLIARI, 1992, p. 30). Nenhum dos trabalhos revisados trataram do papel da inclinação espectral em veicular o acento de palavra em PB (ver sistematização na Tabela 6).

### 2.2.3.2 Processos fonológicos que envolvem o acento

A seguir são apresentados fenômenos que ilustram alguns padrões fonológicos da língua, especificamente aqueles que permitem avaliar implicações teóricas sobre a natureza do acento e sua relação com a prosódia.

Em PB, Major (1992) identificou que as sílabas pós-tônicas sofrem maior encurtamento do que as pretônicas, sugerindo que elas são marcadas por um padrão de duração típico de línguas de ritmo acentual, enquanto as pretônicas seguem um padrão de ritmo silábico<sup>14</sup>. Major propõe que essa diferença de duração tem relação com os padrões distintos de redução vocálica no PB em contextos pretônicos e postônicos. No português europeu (PE), por outro lado, a redução vocálica é mais extrema, afetando inclusive as vogais pretônicas, resultando frequentemente no desvozeamento ou até na elisão da vogal. Esse efeito faz com que o inventário vocálico do PB varie de acordo com a tonicidade da sílaba, i.e. sua posição na palavra em relação à sílaba tônica (WETZELS, 1992; BISOL, 2010b) (ver Figura 7).

A presença de nasalidade em uma vogal parece depender da tonicidade daquela sílaba. Deschamps (1976) afirma que a nasalização só ocorre em vogais tônicas. Por exemplo, em “ama”,

<sup>14</sup> Ver discussão e revisão sobre ritmo acentual ou silábico, proposto por Pike (1945), em Harris (2015).

Contexto Tônico	Vogais Possíveis
Tônica	/i/, /u/, /e/, /o/, /ɛ/, /ɔ/, /a/
Pretônica	/i/, /u/, /e/, /o/, /a/
Postônica Não-final	/i/, /u/, /e/, /a/
Postônica Final	/i/, /u/, /a/

Tabela 7 – Vogais possíveis em cada contexto tônico do Português Brasileiro.

a nasalização está presente na vogal tônica, mas quando essa vogal passa para uma posição átona, como em “amamos”, a nasalidade é perdida.

[Abaurre e Pagotto \(1996\)](#) argumentam que a nasalização nas vogais tônicas é aplicada de maneira categórica, independentemente das características acústicas das vogais, enquanto a variação ocorre apenas quando a vogal alvo da nasalização é átona. Da mesma forma, [Wetzels \(1997\)](#) propõe que a nasalização não distintiva se aplica quase obrigatoriamente às vogais acentuadas e opcionalmente às átonas.

O processo de sândi<sup>15</sup> externo também é sensível à tonicidade das sílabas envolvidas. No caso do PB, a maneira como o sândi vocálico externo atua leva em consideração da tonicidade das vogais justapostas, que devem ser, na grande maioria das vezes, átonas (uma postônica final + uma pretônica inicial) para que o fenômeno ocorra (ver Quadro 8).

Processo	Descrição	Exemplo
<b>Degeminação</b>	Apagamento de uma das vogais iguais entre duas vogais.	meNIna aLÉgre → meni[na]legre
<b>Elisão</b>	Apagamento da primeira vogal entre vogais diferentes.	gaROta eleGANte → garo[te]legante
<b>Ditongação</b>	Reanálise da distribuição dos segmentos nas sílabas.	LEque aZUL → le[kja]zul

Tabela 8 – Resumo dos processos de sândi vocálico externo no português brasileiro.

A ditongação, por sua vez, pode ocorrer ainda que a segunda vogal seja acentuada, desde que a vogal alta na sequência de VV esteja desacentuada ([BISOL, 2002](#)), e.g. “meNIno ALto” torna-se “meNInwALto” ([u] > [w]).

<sup>15</sup> Sândi é um termo da fonologia que se refere a várias mudanças sonoras que ocorrem nas fronteiras entre morfemas ou palavras, por meio da fusão de segmentos ou da alteração de um segmento com base nos adjacentes.

## 2.3 A Fonologia Articulatória

### 2.3.1 Gestos como unidades

A presente dissertação se baseará na proposta da Fonologia Articulatória (FA) (BROWMAN; GOLDSTEIN, 1986; BROWMAN; GOLDSTEIN, 1988; BROWMAN; GOLDSTEIN, 1989; BROWMAN; GOLDSTEIN, 1990) e em outras pesquisas relacionadas a essa teoria. Esta teoria difere das abordagens tradicionais da fonologia, por se concentrar nas ações articulatórias envolvidas na produção da fala, “*as it tries to explain many phonological phenomena as a result of overlap and competition between articulatory gestures*” (HALL, 2010, p. 818). Browman e Goldstein propõem um modelo dinâmico que representa os gestos articulatórios como eventos dinâmicos, o que pode capturar variações acústicas (especialmente a duração), considerando as mudanças na produção ao longo do tempo. Nesse sentido, a abordagem da FA favorece o estudo e análise de maneira a considerar o contexto fonético-fonológico do segmento analisado, fator que será considerado também na presente pesquisa.

A principal proposta da FA é que as unidades representacionais da fonologia correspondem aos eventos de produção da fala, i.e. os primitivos da análise têm base fonética. Dessa forma, a menor unidade fonológica seria o gesto da estrutura motora do aparelho fonador, então, e.g. não o segmento [u], mas suas variáveis do trato (*tract variables*), neste caso: grau de abertura da mandíbula e de arredondamento dos lábios. Esses gestos são descritos como direções para formar ou liberar constrictões no trato vocal (Figura 17).

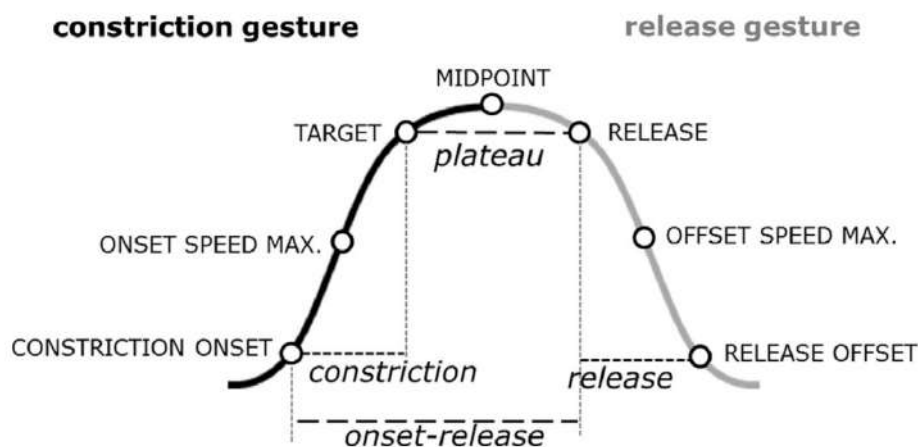


Figura 17 – Pontos e intervalos intragestuais para um gesto típico. Fonte: retirado de Tilsen (2019, p. 28).

### 2.3.2 As dinâmicas de tarefa

De maneira geral, as análises gestuais da FA se baseiam em grande parte em um modelo da física chamado *Task Dynamics* (ou, em português, dinâmicas de tarefa), proposto por Saltzman e Kelso (1987), que se empenha em descrever os movimentos como sistemas dinâmicos orienta-

dos por objetivos específicos, representados por tarefas que guiam o comportamento dos corpos envolvidos. Esse modelo “foi primeiramente desenvolvido para explicar tarefas não linguísticas como atingir um objeto com a mão, ficar de pé ou caminhar” (MARCHAL; REIS, 2012, p. 214). No contexto da fala, isso inclui os padrões temporais de movimentos dos articuladores (e.g. ponta da língua, véu palatino) para produzir sons.

Enraizada na abordagem de sistemas dinâmicos, este modelo postula que os gestos de fala, representados por gestos coordenados, exibem padrões temporais intrincados governados por quatro tipos de propriedades gestuais: intragestural, transgestural, global e intergestural (SALTZMAN; BYRD, 2000). No geral, a dinâmica da tarefa na teoria da FA fornece uma estrutura abrangente para a compreensão da coordenação e do tempo dos movimentos articulatorios subjacentes à produção da fala (KELSO; DEGUZMAN, 1992).

### 2.3.3 A sílaba e o efeito Centro-C

A noção de sílaba continua sendo importante dentro da FA porque essa estrutura oferece uma maneira de entender como os gestos são temporizados e coordenados durante a produção da fala. Essa organização silábica afeta diretamente o comportamento dos gestos articulatorios e explica diversas assimetrias fonológicas observadas entre ataques e codas.

Para a FA, há um foco particular em como o *timing* dos gestos articulatorios varia conforme eles aparecem em diferentes partes da sílaba. Estudos mostraram que os sons no ataque da sílaba, por exemplo, exibem o efeito Centro-C (*c-center effect*) (BROWMAN; GOLDSTEIN, 1988; HONOROF; BROWMAN, 1995; BYRD, 1995), que ajusta a temporização dos gestos em ataque silábico (Figura 19a) para manter um equilíbrio entre seu ponto médio e o núcleo da sílaba. Albano (2020) aponta que Goldstein e colegas

propõem que, no ataque, os gestos articulatorios seguem o regime dinâmico mais comum e estável da natureza, a saber: se alinham com a vogal em torno de zero grau — ou seja, iniciam-se praticamente junto com ela. Esse modo, denominado ‘em fase’, explica a preferência das línguas do mundo pela sílaba consoante-vogal (doravante, CV). Já na coda os gestos articulatorios seguem um regime dinâmico um pouco menos comum e bem menos estável, a saber: se alinham com a vogal em torno dos 180 graus. (ALBANO, 2020, p. 75)

Dessa forma, à medida que mais consoantes são adicionadas ao ataque de uma sílaba, seu ponto médio de ancoragem tende a se manter no mesmo lugar. Ao invés de cada consoante ser coordenada de forma individual com a vogal, o Centro-C cria uma coordenação global da sequência consonantal em relação à vogal seguinte. Por exemplo, em palavras do inglês como ‘spots’, o gesto da língua (para iniciar a produção de [s]) é antecipado em relação a ‘sots’, enquanto o fechamento labial (em [p]) é atrasado em relação a ‘pots’, resultando em um centro de gesto médio (Centro-C) (Figura 18). O mesmo efeito não é presente para os sons na coda, os quais tendem a manter relações de *timing* mais estáveis, sem reajustes (Figuras 18 e 19b). Esse

fenômeno, com base em estudos com o modelo de osciladores acoplados (NAM; SALTZMAN, 2003), indica uma relação “em fase” entre as consoantes do ataque e a vogal. Por outro lado, as consoantes em coda parecem estar em uma relação “anti-fase” ao gesto anterior, seja ele referente ao núcleo silábico ou a um outro gesto consonantal em coda.

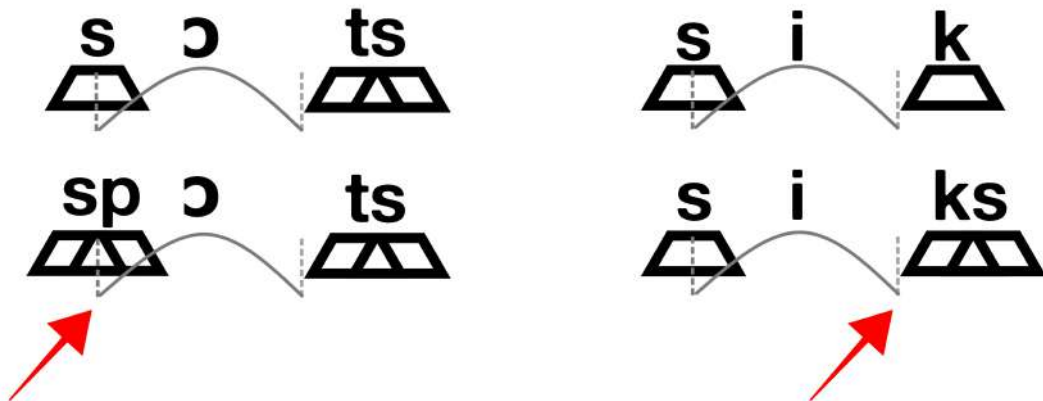
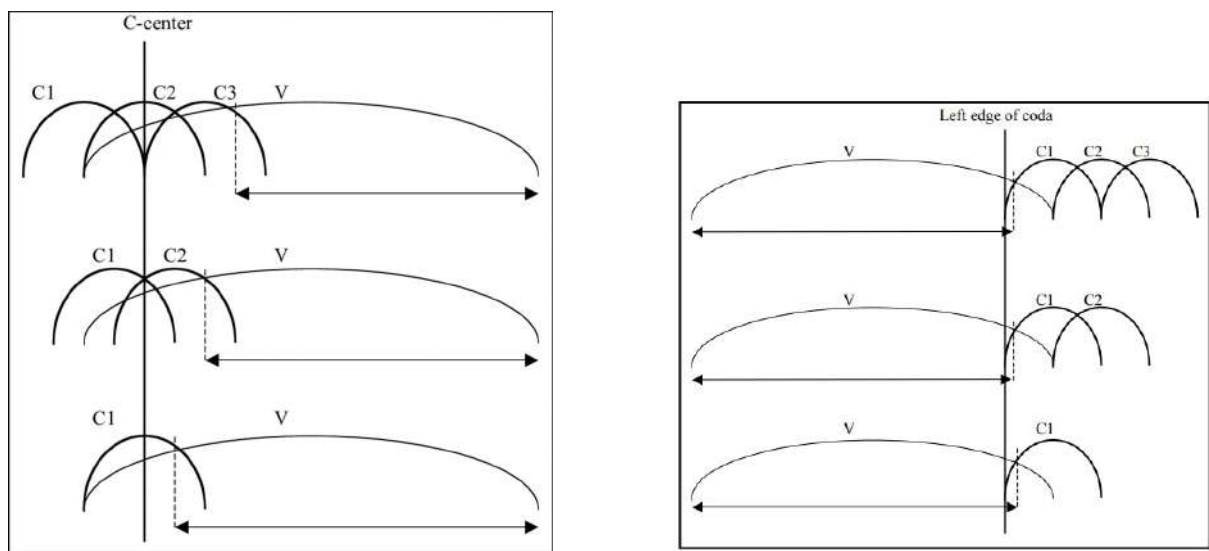


Figura 18 – Efeito de alinhamento de *clusters* consonantais em torno do núcleo silábico em inglês. (a) = *sots* vs. *spots*; (b) = *seek* vs. *seeks*.



(a) Efeito de centro-c.

(b) Efeito de centro-c não presente em codas.

Figura 19 – Comparação do efeito de centro-c em onsets e codas. Fonte: retirado de Katz (2010).

O efeito Centro-C foi observado em algumas línguas além do inglês (TILSEN, 2019, p. 27), como o francês (KUHNERT; HOOLE; MOOSHAMMER, 2006), o italiano (HERMES; MÜCKE; GRICE, 2013), o georgiano (GOLDSTEIN; CHITORAN; SELKIRK, 2007) e o servo (TILSEN et al., 2012). Com existência indicada em outras línguas, esse efeito físico de

organização dos gestos contribui para explicar a preferência de muitas delas por estruturas do tipo CV em vez de VC, uma vez que a relação em fase entre consoante e vogal é mais estável do que a relação fora de fase que ocorre nas codas (HALL, 2010, p. 824).

Embora o efeito centro-c tenha sido amplamente testado em diversas línguas, os resultados são complexos:

*Notably, the C-center effect in sequences of word-initial consonants is not observed in all languages or phonological contexts. Goldstein et al. (2007) found no evidence for a C-center effect in Tashlhiyt Berber: in word-initial consonant sequences, the immediately pre-vocalic consonant was less variably aligned with the vowel than the c-center [...]. Hence the presence of a C-center effect has been taken to be diagnostic of a cross-linguistic typological difference simplex vs. complex organization of word-initial consonants. (TILSEN, 2019, p. 27)*

Estudos confirmam diferenças entre ataques e codas em padrões de sobreposição consonantal, mas nem sempre corroboram as previsões do modelo. Problemas metodológicos, como a variabilidade do ponto de ancoragem usado nas medições, dificultam a interpretação dos dados (TILSEN, 2019, p. 26). Simulações e experimentos sugerem que as medidas temporais isoladas não capturam totalmente as relações de coordenação subjacentes. Além disso, o modelo enfrenta desafios teóricos, incluindo a necessidade de integrar melhor a distinção entre consoantes e vogais e de considerar aspectos como a soância, indicando que o modelo atual requer refinamentos para uma descrição mais completa da dinâmica silábica.

### 2.3.4 A proeminência e o modelamento prosódico na FA

A abordagem de “acento” que pode ser vista na FA até hoje na literatura está mais relacionada a sílabas proeminentes no geral, especialmente ao nível da frase, e não ao da palavra. O trabalho de Iskarous e Pouplier (2022) revisa criticamente como a Fonologia Articulatória aborda os aspectos espaço-temporais da fala, com destaque para a interação entre proeminência e fronteira prosódica, que influenciam sistematicamente os aspectos espaço-temporais dos gestos articulatorios.

Beckman, Edwards e Fletcher (1992), em um estudo avaliando estrutura prosódica e ritmo, mostram que proeminências estão associadas à maior amplitude e duração de gestos, com destaque para a aceleração no gesto de abertura em sílabas acentuadas. Quanto à organização prosódica, fronteiras resultam em aumento de duração, mas afetam desigualmente os gestos: gestos de fechamento (como em /ap#/) sofrem maior impacto em comparação com gestos de abertura (como em /#ap/). Esses fenômenos refletem mudanças nos parâmetros dinâmicos, como rigidez (*stiffness*), mas também sugerem a influência de mecanismos prosódicos externos, como o  $\pi$ -gesture, que explica a redução de sobreposição entre gestos próximos às fronteiras.

Os  $\pi$ -gestures (ou gestos- $\pi$ ), uma adição de Byrd e Saltzman (1998) à FA, são uma proposta teórica para o modelamento da prosódia. Nesse sentido, enquanto os gestos “normais”

(como os de consoantes e vogais) coordenam os movimentos para produzir sons específicos, os gestos- $\pi$  regulam a sincronização temporal e o ritmo de estruturas prosódicas maiores. Esses gestos atuam de maneira transgestual para desacelerar todos os gestos de constrição ativos simultaneamente, em proporção ao nível de ativação do gesto- $\pi$ . Assim como os gestos articulatórios têm propriedades duracionais inerentes e podem se sobrepor a outros gestos, os gestos- $\pi$  também têm uma extensão no tempo e se sobrepõem aos gestos de constrição do trato vocal. Byrd e Saltzman (2003) argumentam, por exemplo, que somente ajustes na rigidez não explicam totalmente as reduções de sobreposição gestual constantemente observadas na fala. Estudos até o momento apontam que esses gestos estão relacionados à organização de acentuação, entoação e fronteiras prosódicas.

Um estudo mais recente, de Roessig e Mücke (2019), ampliou a análise, demonstrando que proeminência afeta simultaneamente contornos tonais e a cinemática articulatória, reforçando a ideia de que os efeitos prosódicos são multidimensionais, influenciando também a articulação gestual. Este estudo apresentou dados sobre a marcação prosódica de foco em alemão, envolvendo 27 falantes. Os seus resultados para entoação indicaram que não há mapeamento direto entre tipos de foco e tipos de acento, mas há tendências probabilísticas que formam padrões estáveis entre quase-categorias. Para a articulação, foi encontrado que proeminências prosódicas são expressas de forma gradual, com ajustes articulatórios, como aumento da abertura labial e rebaixamento da língua na produção da vogal /a/, ampliando sua sonoridade e reforçando traços articulatórios. Estudos como esse possibilitam uma modelagem integrada entre prosódia e articulação, estabelecendo conexões entre estruturas hierárquicas da prosódia e padrões dinâmicos de fala, como o alongamento articulatório e a redução da sobreposição de gestos.

De maneira semelhante, Edwards, Beckman e Fletcher (1991) e Beckman, Edwards e Fletcher (1992) abordam a cinemática do acento com foco nos gestos articulatórios em sílabas acentuadas e não acentuadas. São apresentadas evidências empíricas, utilizando durações médias, deslocamentos e velocidades máximas para gestos de abertura e fechamento, analisados por diferentes velocidades de fala (rápido, normal e lento). Os autores demonstram que as sílabas acentuadas são mais longas do que as não acentuadas, e que esse aumento de duração afeta tanto os gestos de abertura quanto os de fechamento. Quanto a deslocamento no espaço, os gestos articulatórios em sílabas acentuadas são maiores, com uma velocidade do gesto de abertura maior, mas não para o gesto de fechamento. As diferenças nas medidas cinemáticas não são uniformes e indicam que o acento não é realizado como uma modificação uniforme em toda a sílaba. Porém, quanto a isso, mais estudos são necessários em várias línguas, já que

*[w]ithout specific modifications of the profiles of  $\pi$ -gestures, this approach predicts **uniform compression in onset and codas**. It is possible to arbitrarily shape  $\pi$ -gestures in order to mimic local variations in rate, but there is no motivation for doing this in association with onsets or codas specifically. (TILSEN, 2019, p. 30, grifos nossos)*

### 2.3.5 A Fonologia Acústico-Articulatória: proposta de Albano

Albano (2001) aborda a Fonologia Articulatória de Browman e Goldstein com um olhar crítico, reconhecendo seu potencial inovador, mas apontando desafios, principalmente no que diz respeito ao uso de dados estritamente articulatórios em suas abordagens. De acordo com a autora, a teoria se posiciona como uma tentativa de superar as limitações das abordagens simbólicas tradicionais, oferecendo uma perspectiva mais integrada da relação entre o contínuo e o discreto na fala (ALBANO, 2020).

Por um lado, a autora destaca os méritos da FA em propor uma representação temporal que vai além da sucessão linear de unidades, incorporando, por exemplo, a sobreposição dos sons na fala, permitindo uma descrição mais detalhada da variação alofônica e da complexidade segmental. Por outro lado, ela critica o modelo por depender de metáforas físicas, como o sistema massa-mola, e a dificuldade de justificar certos alinhamentos e sobreposições de gestos em termos linguísticos, fisiológicos ou mecânicos:

Browman e Goldstein (1992) lançaram mão de um sistema massa-mola com oscilações altamente amortecidas a fim de modelar o movimento representado por cada caixa. [...]

O problema é que o alinhamento das caixas é imprevisível [...]. Então, ao justapô-las na sucessão, estamos representando um contraste sintagmático da língua, exatamente como fizeram os estruturalistas — o que é natural no caso de uma sequência de fato. No entanto, ao defasá-las e sobrepô-las, estamos propondo um novo tipo de alinhamento que ainda clama por explicação, seja ela linguística, mecânica, fisiológica, comportamental ou outra. (ALBANO, 2020, p. 87)

Ademais, a autora enfatiza a necessidade de incorporar aspectos sensoriais e acústico-articulatórios ao modelo, sugerindo que a especificação puramente motora dos gestos não é suficiente para dar conta da complexidade da fala. Sua crítica se estende à falta de integração com teorias mais amplas de controle motor, particularmente no que diz respeito ao papel do *feedback*, um componente essencial que está ausente da implementação tradicional da fonologia gestual.

Nesse contexto, Albano propõe uma reorientação da FA, ampliando sua base epistemológica e metodológica para incluir tanto as dimensões sensório-motoras quanto as cognitivas e comportamentais da fala. Ao mesmo tempo, ela reconhece que o conhecimento atual ainda é limitado, especialmente no que tange às unidades superiores, como a sílaba e a palavra, e à organização interna das relações entre os gestos.

O presente capítulo abordou de maneira detalhada a Fonologia Articulatória, concentrando-se nas propostas de Browman e Goldstein e em outros estudos relacionados. A análise dos gestos articulatórios como unidades fundamentais na produção da fala revela como esses movimentos se coordenam de maneira complexa, sendo governados por dinâmicas de tarefa que organizam o tempo e a sobreposição entre os gestos. A importância da estrutura silábica, especialmente no que

se refere ao efeito centro-c, também foi destacada, mostrando como a coordenação dos gestos varia dependendo da posição silábica, com implicações importantes para a compreensão da estrutura fonológica das línguas. Foi possível observar que a relação entre sílaba, proeminência e o modelo dinâmico da FA oferece um quadro teórico e metodológico coerente para a investigação do acento de palavra. Com base no embasamento teórico apresentado, o próximo capítulo apresenta a metodologia empregada no experimento, detalhando os materiais, os participantes e os procedimentos adotados para a coleta e o tratamento dos dados.

### 3 Metodologia

O fluxo das atividades realizadas para a realização da presente pesquisa está na Figura 20. A descrição, discussão e justificativa dessas atividades será feita nas seções a seguir. Primeiramente, será apresentado o experimento realizado para a obtenção do material a ser analisado, depois serão apresentadas as etapas de anotação, revisão, extração e análise dos dados obtidos.

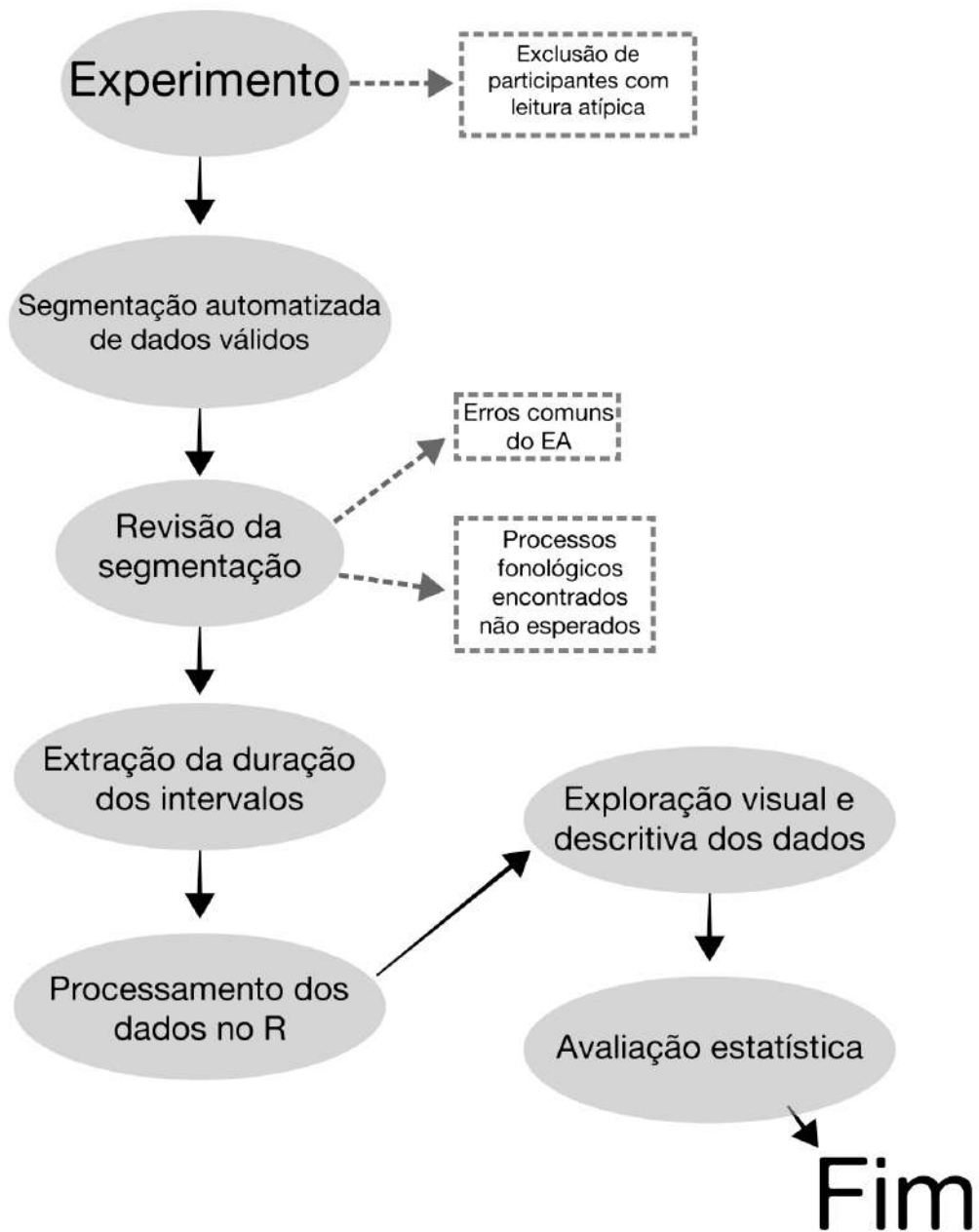


Figura 20 – Fluxograma dos processos referentes à metodologia deste trabalho.

## 3.1 Experimento

### 3.1.1 Materiais

Para medir sílabas e segmentos em diferentes contextos acentuais (acentuado/não-acentuado), um experimento foi desenhado para se obter os dados necessários para as medições e análises.

As palavras-alvo foram selecionadas de modo que se tivessem, em contextos tônicos e pós-tônicos, sílabas contendo 14 consoantes diferentes no ataque e, ainda, em quatro moldes silábicos distintos (CV, CCV, CVC e CCVC) ( $n = 282^1$  palavras x  $14^2$  participantes = 3384 palavras analisáveis;  $n$  final de segmentos medidos: 6561). Para cada consoante em cada molde silábico, três palavras foram selecionadas, pensando em diluir o efeito do item lexical para a análise de dados, por exemplo: para sílabas do tipo CV iniciadas por [ʒ] tem-se ‘loja’, ‘cuja’ e ‘fuja’ para o contexto pós-tônico e ‘cajá’, ‘Iemanjá’ e ‘guarujá’ para o tônico; ou para sílabas do tipo CVC iniciadas por [n], ‘dunas’, ‘runas’, ‘zonas’, ‘guaranás’, ‘tenaz’ e ‘zonás’<sup>3</sup> (ver Tabela 9).

Molde silábico	Consoante	Palavra-veículo (sílabas átona)	Palavra-veículo (sílabas tônicas)
CV	g	manga	ingá
CV	k	vaca	picá
CVC	g	cegas	ingás
CVC	k	vacas	tacacás
CCV	g	magra	migrá
CCV	k	lucra	lacrá
CCVC	g	magras	migrás
CCVC	k	lacrás	lacrás
CV	v	luva	mandruvá
CV	f	fofa	sofá
CVC	v	luvas	mandruvás
CVC	f	fofas	sofás
CCV	v	palavra	lavrá
CCV	f	saфра	chifrá
CCVC	v	palavras	lavrás
CCVC	f	safras	chifrás

Tabela 9 – Exemplos de palavras utilizadas no experimento com as oclusivas velares e as fricativas labiodentais. Conjunto de dados completo em Anexo B.

<sup>1</sup> Este número corresponde às palavras originais (264) somadas àquelas 18 (apresentadas posteriormente na Tabela 15) adicionadas pós-piloto para averiguar o efeito de fronteira de palavra na coda da sílaba. Ver discussão sobre isso na Subseção 3.2.5.

<sup>2</sup> 18 menos quatro que precisaram ser excluídos, conforme será explicado na Subseção 3.2.1

<sup>3</sup> Para garantir um desenho experimental completo e comparável, foram adicionados logotomas ao conjunto de dados. Em alguns casos, nomes próprios também foram adicionados. Dentre as 282 palavras do experimento, cerca de 48 (aproximadamente 17%) eram logotomas.

A coleta foi realizada na Faculdade de Letras da UFMG, no quarto andar, dentro de uma cabine acusticamente isolada, garantindo qualidade de gravação adequada. O participante e o pesquisador permaneciam na sala durante todo o experimento, com o pesquisador controlando o software e orientando o participante quando necessário.

Para a gravação, foi utilizado um microfone de mesa fixo (pertencente ao Laboratório de Fonologia da Faculdade de Letras) e um computador pessoal. A gravação foi realizada através do software Audacity (AUDACITY Team, 2023), conforme mostrado na Tabela 10.

Item	Configurações
Gravador	Zoom H4n
Áudio	Mono
Formato	WAV
Taxa de amostragem	44100Hz
Modo de conexão	Conectado ao USB de um computador
Software utilizado	Audacity (AUDACITY Team, 2023)
Distância do participante	Aproximadamente 40cm

Tabela 10 – Configurações do equipamento e software utilizados na coleta de dados.

### 3.1.2 Procedimentos de coleta

Para obter os dados necessários para o estudo, áudios de falantes nativos do PB foram gravados e analisados. A coleta foi realizada em um experimento de produção no qual os participantes foram instruídos a ler frases-veículo (“Eu digo [...] baixinho.”) que continham as palavras-alvo<sup>4</sup>. Também foi considerado um possível efeito confundidor que seria a resolução de um choque acentual entre a última sílaba da palavra-alvo e a primeira do advérbio que seguia.

Cada participante foi informado sobre os detalhes do experimento, incluindo a duração aproximada de 10 a 15 minutos, dependendo da velocidade de fala. Eles foram instruídos de que leriam cerca de 200 palavras inseridas em frases, e que algumas dessas palavras eram *logatomas* — palavras inexistentes no português, mas “pronunciáveis”. Também, foi pedido a eles que lessem as frases da maneira “o mais natural possível”, na intenção de não fazê-los incidir foco sobre nenhuma palavra, especificamente a alvo.

Além disso, um *script* em R (Anexo C) foi utilizado para gerar uma sequência de apresentação de slides (em formato PowerPoint) diferente para cada participante, garantindo que a ordem de apresentação das frases fosse aleatória, a fim de evitar qualquer efeito de ordem no futuro modelo estatístico.

<sup>4</sup> Nas fases iniciais do experimento (piloto), a coleta foi realizada com cinco participantes, utilizando um corpus reduzido de 264 palavras-alvo, excluindo aquelas que avaliavam codas mediais. Além disso, a frase-veículo utilizada era “Eu digo [...] toda hora.”, mas foi posteriormente substituída por “Eu digo [...] baixinho.” devido à conclusão de que o [t] que seguia a palavra-alvo poderia interferir na produção do segmento [s] em sílabas com codas preenchidas, por conta do ponto de articulação compartilhado. Procurou-se, também, evitar algum possível tipo de efeito de choque acentual causado pelo encontro de uma palavra oxítônica com a primeira sílaba tônica de ‘toda’.

Antes de iniciar a gravação, três slides de treinamento, contendo as mesmas palavras (que não faziam parte do conjunto analisado) para todos os participantes, foram exibidos para assegurar que o participante entendesse como deveria proceder. O experimento era autogerido: os participantes eram responsáveis por avançar os slides no seu próprio ritmo, utilizando a seta do teclado do computador.

Se um erro fosse cometido durante a leitura de uma palavra, os participantes eram instruídos a simplesmente repetir a palavra. Se o erro não fosse percebido pelo participante, o pesquisador interviria e solicitava a repetição da palavra.

### 3.1.3 Participantes

A amostra do experimento consistiu em 18 participantes. Todos os participantes tinham entre 18 e 32 anos e estavam matriculados em cursos de graduação na UFMG. A maioria dos participantes era oriunda do Complexo da Faculdade de Filosofia e Ciências Humanas (FAFICH), com alguns participantes da Escola de Música da UFMG. Todos os participantes possuíam o português brasileiro como língua materna, e cumpriam os requisitos especificados no Termo de Consentimento Livre e Esclarecido (Anexo A).

Conforme mencionado anteriormente, os áudios resultaram em um total de 6531 fones analisáveis. Caso o número de participantes gravados (importante para o efeito aleatório de cada indivíduo no modelo linear misto) ou de segmentos analisados não fosse suficiente para garantir bons ajustes estatísticos, novas sessões de gravação seriam organizadas com outros participantes.

## 3.2 Tratamento dos dados

### 3.2.1 Dados desconsiderados

Embora 18 pessoas tenham sido gravadas, nem todos os 18 áudios foram utilizados para a análise do objeto de estudo. Houve casos em que o participante teve dificuldade de ler corretamente as frases (participantes 8, 12, 13 e 18), o que resultou em uma cadência de leitura hesitante, i.e. com muitas pausas espalhadas entre as palavras ao longo das frases<sup>5</sup>.

Paralelamente, algumas frases de alguns participantes não foram analisadas acusticamente devido a diversos fatores, como: pausa(s) (e outros tipos de hesitações ou emissões sonoras não-linguísticas, como tosses ou espirros) entre palavras da mesma frase<sup>6</sup>; e/ou leitura incorreta de palavras<sup>7</sup>.

<sup>5</sup> Os participantes de número 12 e 13, especificamente, realizaram pausas após palavras-alvo oxítonas quase categoricamente.

<sup>6</sup> O desenho de dados e de experimento considera que todas as palavras devem estar na mesma posição prosódica dentro da frase. No caso do presente trabalho, os sons analisados devem estar em palavras em posição medial, ou seja, sem pausas em suas fronteiras.

<sup>7</sup> Em alguns casos, os participantes não leram as palavras considerando o acento ortográfico, o que resultaria na produção de um item diferente daquele considerado para o desenho do experimento.

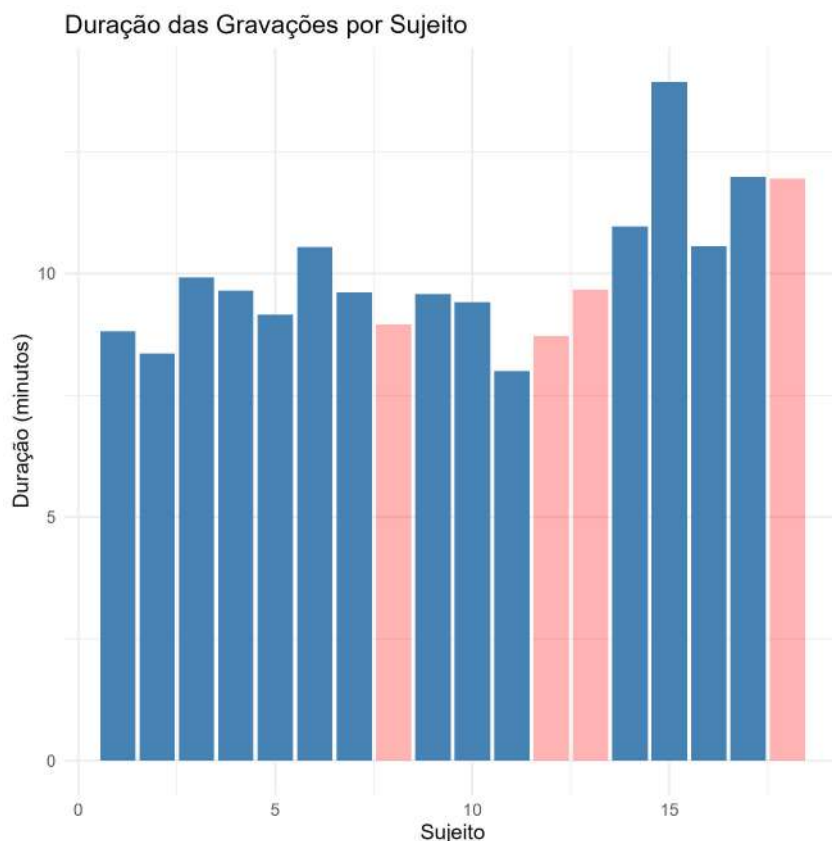


Figura 21 – Duração dos áudios gravados de cada sujeito. Estão em rosa os participantes excluídos da análise.

### 3.2.2 Manipulação e etiquetagem dos áudios gravados

O grande número de dados utilizado no presente trabalho envolve uma alta carga de trabalho, principalmente de segmentação e anotação, que pôde ser amenizada com o uso de ferramentas automatizadas, como *scripts* em R e em Praat e *softwares* úteis e convenientes. Um dos programas utilizados foi o EasyAlign (EA) (GOLDMAN, 2011), que é uma ferramenta de alinhamento fonético automático baseada no *toolkit* HTK (YOUNG et al., 2002), integrada ao Praat (BOERSMA; WEENINK, 2023) e com suporte a várias línguas, como o francês, o espanhol e o PB.

O EA gera um TextGrid com múltiplas camadas: de fones, sílabas, palavras e enunciados, a partir de um áudio e sua transcrição<sup>8</sup>. O processo envolve três etapas automáticas principais: macrosegmentação, conversão grafema-fonema e segmentação em fonemas, entre as quais podem ser necessários ajustes manuais<sup>9</sup>. As camadas resultantes são: ‘*phones*’ (fones), ‘*syll*’ (sílabas), ‘*words*’<sup>10</sup> (palavras), ‘*main\_words*’ (palavras-alvo), ‘*phono*’ (transcrição fonética) e

<sup>8</sup> O *script* utilizado para gerar o *input* do experimento de produção (Anexo C) também gera um arquivo .txt com as frases separadas por quebra de linha, como necessário para a primeira etapa do EA.

<sup>9</sup> Após à etapa de conversão grafema-fonema, foi removido da recém-criada camada de fonemas o caractere “j” inserido pelo EA em certos contextos. Para essa tarefa, foi utilizado o *script* “*remove\_j.praat*” (Anexo D).

<sup>10</sup> O *tier\_cleaning.praat* (Anexo E) foi rodado para duplicar a camada de palavras e excluir os rótulos que continham palavras que não fossem as alvo, i.e. “eu”, “digo” e “baixinho”.

‘ortho’ (transcrição ortográfica). O resultado é um arquivo TextGrid com várias camadas de intervalo, conforme a Figura 22.

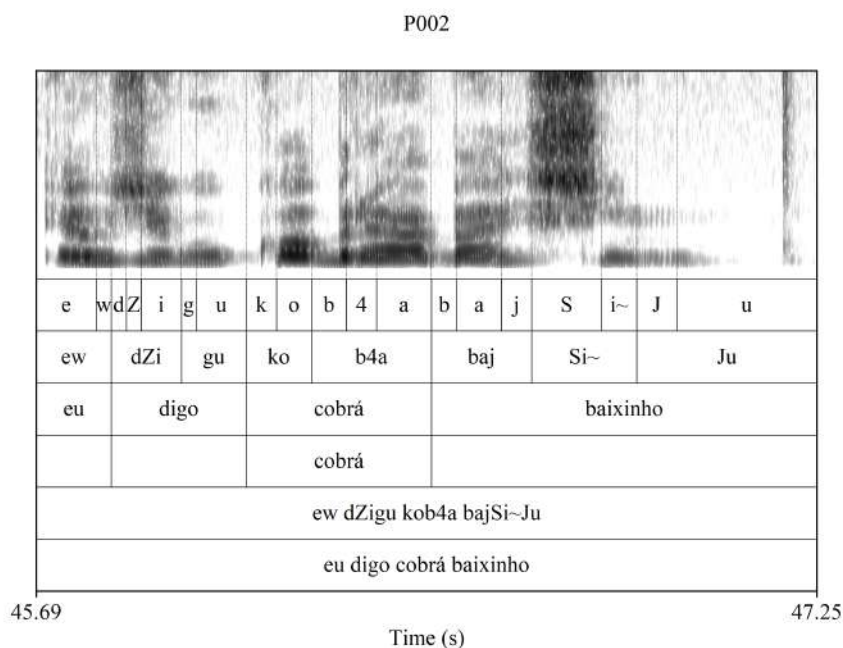


Figura 22 – Trecho de um TextGrid recém-gerado pelo EasyAlign. Legenda: 1 - Camada “*phones*”; 2 - “*syll*”; 3 - “*words*”; 4 - “*phono*”; 5 - “*ortho*”.

Finalmente, finalizado o trabalho automático de programas e *scripts*, foi iniciada a etapa de revisão dos arquivos TextGrid. Todas as últimas sílabas, que seriam as porções das palavras a serem analisadas posteriormente, das palavras-alvo tiveram suas segmentações revistas pelo pesquisador, e uma atenção especial foi dada a casos que pareceram mais desafiadores para o trabalho do EA. O método utilizado para corrigir erros específicos e mais problemáticos do EA teve dois critérios: um de procurar ativamente por segmentos que se demonstraram mais suscetíveis a erros ([r]<sup>11</sup>, [s]<sup>12</sup> e a maioria das consoantes oclusivas<sup>13</sup>); e, posteriormente, na fase exploratória de dados, os segmentos com valores atípicos (*outliers*) de duração tiveram suas segmentações revisitadas e conferidas novamente.

### 3.2.3 Fenômenos fonológicos não esperados

A seguir serão apresentados outros casos que demandaram a revisão da segmentação e anotação dos TextGrids, não por causa do EasyAlign, mas sim por fenômenos fonológicos que

<sup>11</sup> A segmentação do EA atribuía ao tepe uma porção do áudio muito maior do que a real. Este erro do programa foi aquele mais crítico para a futura análise dos dados, pois envolvia a diminuição considerável da duração dos sons que faziam fronteira com o tepe.

<sup>12</sup> A fricativa alveolar desvozeada sofria com a repetição de valores idênticos em sua duração, algo provavelmente relacionado ao método do HTK, em que o EA se baseia para realizar suas funções. O problema maior foi que esses valores poucas vezes correspondiam à real duração dessa consoante, o que foi devidamente ajustado.

<sup>13</sup> As principais características da segmentação automatizada de oclusivas pelo EA são (i) a marcação do início de oclusão muito mais tardio do que o real, partindo do que se vê nos dados acústicos do espectrograma e oscilograma; e (ii) o ponto de soltura consideravelmente mais tardio do que o verificado de fato, tomando parte da vogal para a duração da consoante.

não haviam sido considerados durante a elaboração do desenho experimental.

### 3.2.3.1 Casos de vozeamento do [s] em coda

Outra intervenção que precisou ser feita diz respeito à anotação de [s] realizada pelo EasyAlign, pois esse som, quando em final de palavra, passou por um processo de assimilação de vozeamento com o som seguinte [b], referente à palavra “baixinho” da frase-veículo. A mudança de rotulagem é realizada para tornar possível a separação entre [s] e [z] na futura descrição e demais análises dos dados. Caso tal fenômeno fosse ignorado, a duração dos numerosos segmentos vozeados poderia mascarar alguns efeitos duracionais, e.g. de acento ou de molde silábico, e fazer aparecer resultados irreais ou mascarar aqueles que se procura de fato, já que consoantes vozeadas são mais curtas que seus pares desvozeados (ABRAMSON; LISKER, 1964; KEATING, 1984; DOCHERTY, 1992; KLUENDER; DIEHL; WRIGHT, 1988).

O fenômeno de vozeamento do /s/ em final de sílaba no português brasileiro está associado ao ambiente fonológico seguinte. O /s/ concorda em vozeamento com o contexto: ele é surdo antes de consoantes surdas ou pausas e vozeado antes de vogais ou consoantes vozeadas (ALBANO, 1999, p. 35).

Embora o *plugin* considere alguns tipos de processos fonético-fonológicos<sup>14</sup>, essa alteração de vozeamento passa despercebida, uma vez que ele se baseia em informações ortográficas pré-estabelecidas para fazer a transcrição fonética dos sons. Por isso, foi necessária a elaboração de um *script* em Praat para analisar acusticamente os intervalos marcados como [s] e sucedidos de [b], e alterar seus rótulos para “z”, se for constatado o vozeamento.

Primeiramente, foi preciso decidir qual marca acústica apontaria para a presença de consoante vozeada no intervalo do som. Essa tarefa, simples a princípio, mostrou-se marcada por alguns empecilhos. Um deles, e o principal, é o caráter gradiente do processo fonológico, como bem retrata Albano (1999):

a análise espectrográfica do PB revela uma situação mais complicada. O processo é categórico em posição medial, mas **torna-se gradiente em posição final**, dependendo da força da fronteira seguinte. Fronteiras fracas propiciam o vozeamento total. Fronteiras fortes inibem-no ou tornam-no parcial. (ALBANO, 1999, p. 35, grifos nossos)

Esse cenário está intimamente ligado à solução encontrada para a nova rotulagem de “s” para “z”. No presente trabalho, não será avaliada a situação do fenômeno de vozeamento por assimilação em PB. Dito isso, a fim de viabilizar a automatização do processo de nova rotulagem dos intervalos, foi preciso estabelecer um ponto fixo que separasse um som vozeado de um desvozeado. Entende-se que considerar o vozeamento uma característica gradiente tornaria

<sup>14</sup> Por exemplo: elisão vocálica, que pode ser considerada ou ignorada por meio de uma configuração do próprio usuário; e palatalização de vogais, como em [ha.'pas] > [ha.'paʲs].

a presente pesquisa bem mais complexa, o que faria ser necessário mais tempo para a sua (re)elaboração.

Ainda que se tentasse encontrar um ponto fixo e objetivo para separar a fricativa alveolar vozeada da não vozeada, outro problema foi encontrado, já que as principais medidas em que se pensou automatizar para a tarefa não estavam, após conferência, sendo de fato bons apontadores de vozeamento. A primeira medida foi o a Razão de Harmonicidade para Ruído (*Harmonicity-to-Noise Ratio*, ou simplesmente HNR). Ainda que se esperasse que a periodicidade das ondas de vozeamento fossem elevar o valor de HNR, isso não foi o que aconteceu. Além disso, o programa Praat retornava valores diferentes para a medida de acordo com o método de medição escolhido (pelo menu na janela *Praat Objects* ou na de *Textgrid*), o que não foi possível de ser contornado. Também foi considerado utilizar o *pitch*, pensando que valores válidos corresponderiam à consoante vozeada, e os inválidos, às desvozeadas. Porém, além de o programa Praat ter sido capaz de retornar valores não-nulos de *pitch* para [s], o *script* programado para a tarefa de extrair informações acústicas e ajustar a rotulagem não lidava bem com valores nulos.

A solução encontrada foi não ideal, à medida que se beneficiou de um erro de cálculo do Praat: foram extraídos os valores do primeiro formante (F1); aquelas medidas que fossem menores que 600Hz indicariam um cálculo válido de formantes (por ser um valor esperado) e, portanto, refletiria que o som é vozeado. Ainda assim, a verificação oitiva, pelo próprio pesquisador, do método adotado apontou que houve um alto grau de sucesso em se refletir o vozeamento da fricativa alveolar.

O *script* (Anexo F), portanto, percorre a camada de segmentos, procurando rótulos “s” seguidos de “b”. O intervalo marcado como “s”, então, é separado em quatro partes, e as duas mais ao meio (porções 2/4 e 3/4) são selecionadas<sup>15</sup>; ali se calcula o valor do primeiro formante. Caso o valor indique o vozeamento da consoante, o intervalo ganha um novo rótulo (“z”) e é anotado, em uma nova camada, que essa alteração foi feita. Os novos TextGrids são salvos, por fim, em uma outra pasta.

É importante citar que realizar essa tarefa de maneira automatizada permitiu analisar e corrigir um grande número de arquivos de forma rápida e eficiente, economizando tempo e esforço, já que o número de intervalos a ser avaliado seria em torno de 1.400. Ainda que a nova rotulagem de “s” para “z” fosse feita manualmente, não se escaparia do caráter gradual do vozeamento, que pode ser difícil de julgar auditivamente.

O cenário final, após alterações de rotulagem, ficou como se vê na Tabela 11.

<sup>15</sup> Isso foi feito para a extração de dados acústicos ser feita em uma porção estável da consoante, sem tanto efeito dos sons ao redor.

Tabela 11 – Distribuição das ocorrências de segmentos em posições de ataque e coda.

Segmento	Em posição de ataque	Em posição de coda
[s]	134	449
[z]	124	899

### 3.2.3.2 Epêntese segmental

Constantemente, ao se segmentar e anotar áudios de fala, como neste estudo, emergem fenômenos não previstos relacionados à epêntese, os quais resultam na inserção de segmentos adicionais em contextos específicos. Esses segmentos, embora não previstos no molde silábico original, apresentam implicações para a análise, exigindo decisões cuidadosas sobre sua inclusão ou exclusão da análise. A seguir, exploram-se dois casos distintos de epêntese segmental (especificamente vocálica) encontrados no conjunto de dados desta pesquisa e justificam as abordagens adotadas para lidar com tais fenômenos.

No PB, o tepe [r], como ocorre em palavras como *prata* e *preta*, é uma consoante vibrante simples caracterizada por uma única e breve oclusão da língua nos alvéolos (SILVA, 2003; MATEUS; D'ANDRADE, 2000; SILVEIRA; SEARA, 2017). As consequências disso para o sinal acústico incluem a presença de um espaço em branco no espectrograma (Figura 23), indicando a oclusão momentânea do trato vocal, acompanhada por um declínio da amplitude da onda sonora (MENDES JÚNIOR, 2018, p. 18). O tepe, então, exibe propriedades acústicas que o aproximam de consoantes oclusivas, devido ao bloqueamento rápido do trato vocal, e de semivogais, por apresentar perturbação na estrutura formântica das vogais ao seu redor, com F3 abaixado e próximo a F2 (KENT; READ, 2002).

Pode ocorrer que o tepe, em encontros consonantais, seja precedido por um elemento vocálico (Figuras 24 e 25). Dois moldes silábicos analisados na presente pesquisa (CCV e CCVC) forneciam, então, um ambiente favorável para o aparecimento desse novo som. Para Barbosa e Madureira (2015), a explicação sobre esse cenário de emergência de um segmento vocálico está relacionada a um certo grau de coarticulação entre os sons da sílaba onde o encontro consonantal ocorre:

Esse segmento acústico intercalado entre a oclusiva e a vogal seguinte caracteriza o *tap* em *onset* [(ataque)] complexo em qualquer língua. Isso se dá porque, no momento de soltura da oclusiva, o corpo da língua, por coarticulação antecipatória, já se encontra a caminho da posição da vogal que segue o *tap*, mas a ponta da língua ainda não tocou a região alveolar. Essa configuração da língua produz acusticamente uma vogal, visto que é a configuração articulatória de uma vogal. (BARBOSA; MADUREIRA, 2015, p. 541)

O presente trabalho o chamará de “vogal de apoio” (doravante VA), escolhendo um dentre vários nomes que esse elemento vocálico recebeu ao longo da literatura:

*I utilize the term ‘intrusive vowel’ in the current study. Observe that it is also referred to in the literature as ‘the epenthetic’, ‘epenthetic vowel’, ‘excrescent*

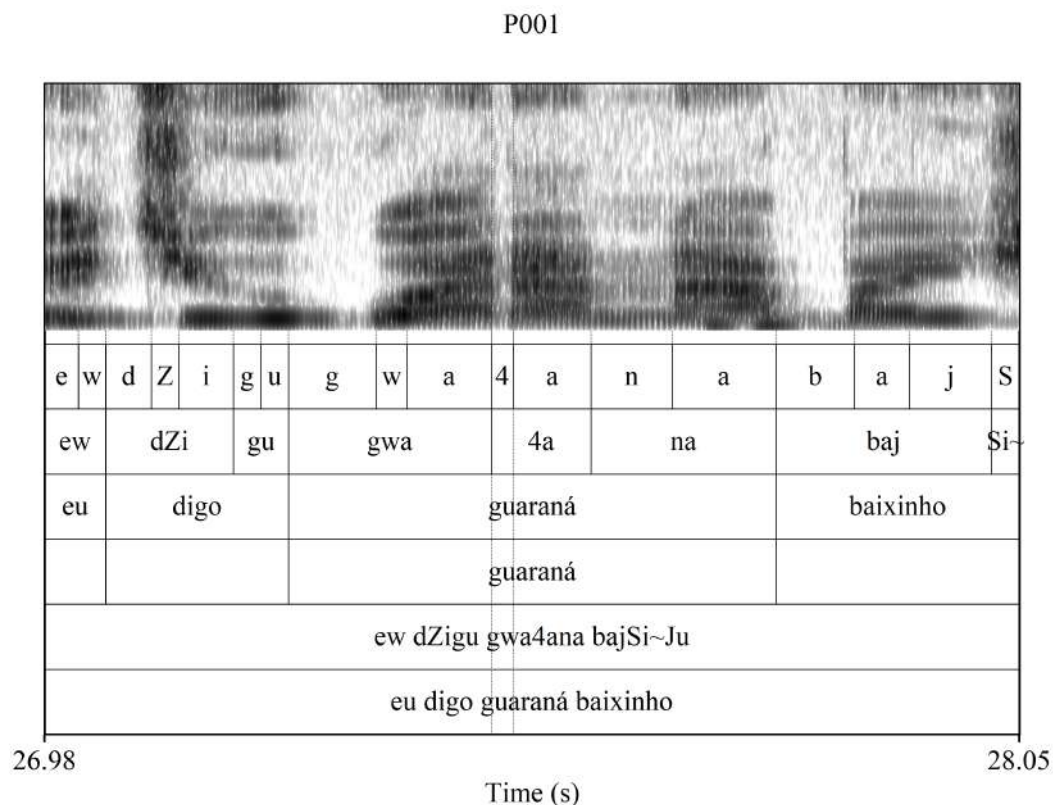


Figura 23 – Exemplo de espectrograma de um tepe intervocálico na palavra ‘guaraná’. “4” = [r].

*vowel*, ‘*schwa insertion*’, ‘*excrescent schwa*’ or ‘*svarabhakti vowel*’. (SCHMEISER, 2009, p. 207, nota 3)

Embora a VA traga complexidade à análise proposta neste trabalho, a segmentação do tepe foi simplificada, medindo o segmento em sua totalidade, sem separar a duração da VA e da obstrução. Tal decisão foi tomada entendendo que isso aumentaria a complexidade desta pesquisa, que, por sua vez, está focada também em outros vários segmentos. Admite-se, ainda, que uma análise que considere as subpartes do tepe é indispensável para estudos que se propõem a descrever o efeito do acento de palavra para esse tipo de som. No entanto, entende-se que o foco do presente trabalho é analisar o efeito do acento de palavra em sílabas de diferentes moldes e em segmentos em diferentes constituintes silábicos, e análise do tepe acontece pela coincidência de ele ser o único segmento que ocorre na segunda posição de ataques complexos no presente conjunto de dados<sup>16</sup>.

As consequências da abordagem escolhida nesta pesquisa serão melhor abordadas na parte de resultados sobre a cauda do ataque complexo (4.2.2.2), mas principalmente em 5.2.2.2.

De forma similar ao cenário encontrado anteriormente, em alguns casos, após palavras de última sílaba CVC ou CCVC, i.e. com coda, foram verificados casos de inserção de um gesto

<sup>16</sup> Caso fosse inserido algum outro segmento com o qual o [r] pudesse ser comparado, o número de palavras a serem lidas pelos participantes tornaria a tarefa de leitura ainda mais cansativa. Por isso, a análise de outros segmentos nessa posição silábica (e na coda, que está aqui preenchida somente por [s] ou [z]) fica a cargo de futuros trabalhos.

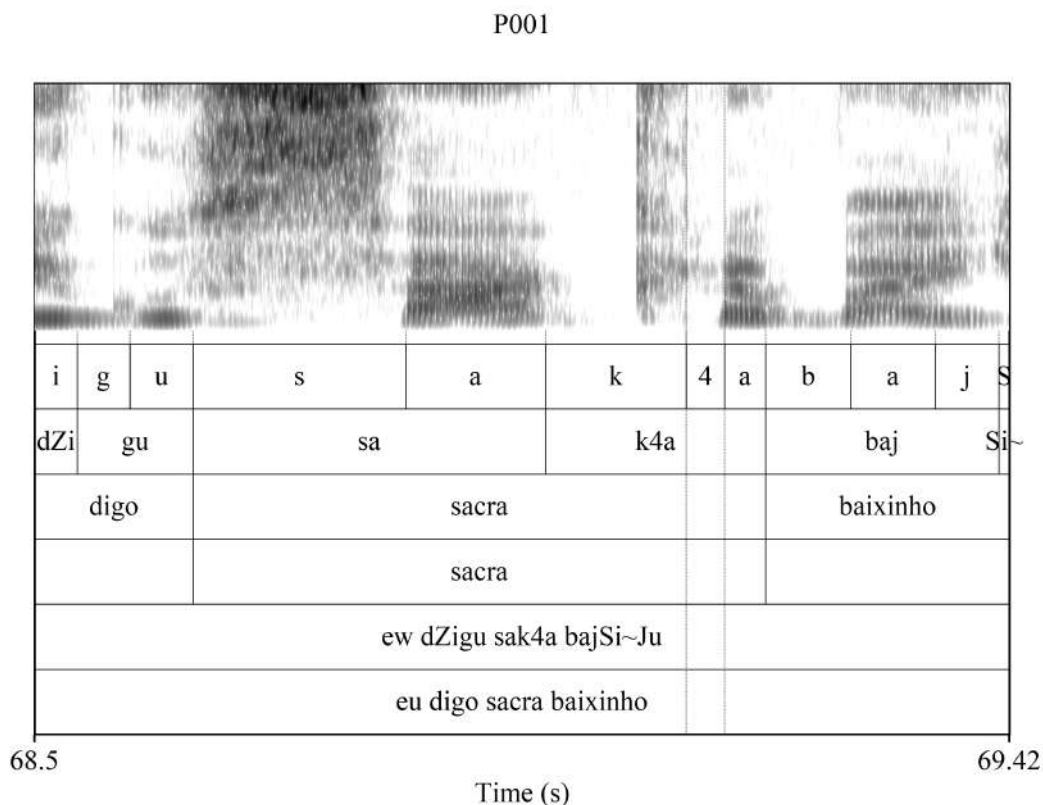


Figura 24 – Exemplo de espectrograma de um tepe em encontro consonantal na palavra ‘sacra’. “4” = [r].

vocálico, potencialmente por conta do vozeamento presente no gesto seguinte da palavra ‘baixinho’<sup>17</sup>. Por exemplo, em “Eu digo pastas baixinho”, algo como [ˈpas.tɐ.sə] foi ocasionalmente produzido (ver Figura 26).

A abordagem escolhida foi a de não analisar esse segmento vocálico epentético, deixando sua duração de fora da sílaba analisada para o presente estudo. A inclusão de uma vogal epentética implicaria a análise de duas sílabas consecutivas com o molde CV, o que desbalancearia o desenho experimental originalmente proposto e aumentaria a complexidade dos métodos de análise do trabalho. Como este estudo não previa a análise de sílabas CV consecutivas nesses contextos, optou-se por manter a classificação das sílabas como CVC e excluir a duração da vogal epentética da análise.

Cabe ressaltar que a ocorrência desse fenômeno não foi categórica para todas as sílabas CVC. Também, alguns participantes realizaram um maior número de epênteses, enquanto outros não realizaram, aparentemente, uma vez sequer. Por esses motivos, entende-se que a modelagem estatística pode dar conta de amenizar os efeitos da vogal epentética na duração de /s/ em coda, por considerar como efeito aleatório a fala individual de cada sujeito (cf. 4.2.2.3).

Após a revisão geral e específica dos arquivos, os TextGrids estavam prontos para terem

<sup>17</sup> Ver a Subseção 3.2.3.1 para discussão sobre o vozeamento de [s] em coda final, também causado por assimilação do processo de vozeamento de [b].

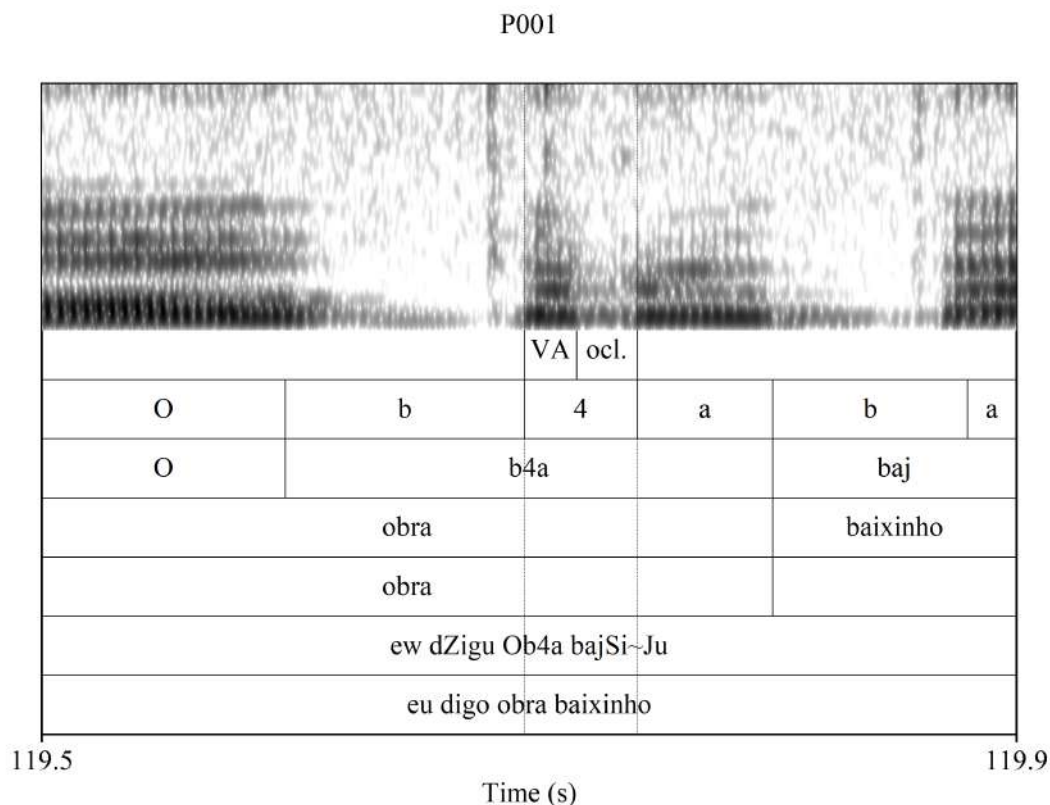


Figura 25 – Exemplo de espectrograma de um tepe com vogal de apoio em encontro consonantal na palavra ‘obra’. ‘4’ = [r].

as durações de segmentos extraídas e analisadas, conforme mostra a Figura 27.

### 3.2.4 Processamento de Dados no R

O processamento dos dados foi realizado no software R (R Core Team, 2021) a partir de arquivos *.txt*, utilizando os dados extraídos com o *extract.praat*<sup>18</sup>. O R foi utilizado para que fossem geradas novas variáveis a partir do produto do *script* do Praat. Um extenso *script* em R foi programado para somar, às seis colunas originais (*part*, *ortholabel*, *seglabel*, *sylllabel*, *wordlabel*, *segdur*, *sylldur* e *worddur*), outras 16 novas. Todo o processo está descrito a seguir.

Após a importação, as colunas relevantes foram convertidas para tipos numéricos ou categóricos, enquanto se corrigiam inconsistências nos rótulos de palavras. Nesse momento também foram excluídas do conjunto de dados aquelas palavras irrelevantes, pois faziam parte da estrutura padrão da frase-veículo: ‘eu’, ‘digo’ e ‘baixinho’. Também removemos duas palavras problemáticas, ‘ingá’ e ‘desfaz’, devido a dificuldades na separação de sequências consonantais ([ng] e [sf], respectivamente), o que poderia comprometer as análises.

<sup>18</sup> O *script extract.praat*, cujo código completo está no Anexo G, realiza a leitura de arquivos no formato *.TextGrid* localizados em uma pasta predefinida e extrai informações a duração em segundos de segmentos, sílabas, palavras e frases anotadas nos camadas correspondentes.

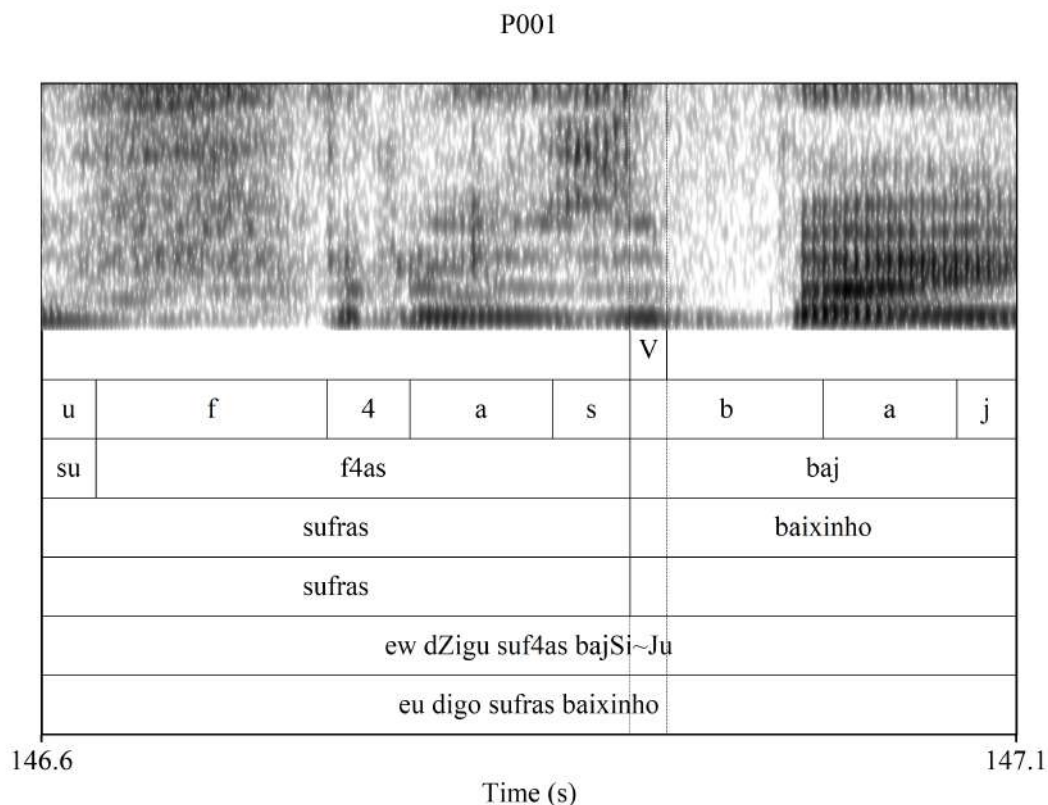


Figura 26 – Espectrograma do contexto fonológico do logotoma ‘sufRAS’ com o fenômeno de epêntese vocálica pós-/s/.

Foram criadas diversas variáveis auxiliares essenciais para a análise, incluindo: tipo do segmento (*segtype*), que categoriza segmentos como vogais ou consoantes; tonicidade da palavra (*wordstresstype*), que identifica palavras oxítonas ou paroxítonas; número de sílabas e segmentos; posição do segmento na sílaba (*segsyllpos*) e da sílaba na palavra (*syllpos*); tonicidade da sílaba (*stress*), categorizando-a como tônica ou átona; estrutura silábica (*syllstruc*), preenchida como CV, CVC, CCV ou CCVC. Essas variáveis foram geradas utilizando diversas funções do pacote *dplyr* e expressões regulares.

Adicionalmente, foram criadas variáveis relacionadas à codificação hierárquica da sílaba (*sylltree*), distinguindo ataques simples (*ons*), ramificados (*ons1* ou *ons2*), núcleos (N) e codas (*cd*), além de características articulatórias para as consoantes, como vozeamento (*voz*), ponto de articulação (*artp*) e modo de articulação (*artm*).

Por fim, algumas variáveis normalizadas foram criadas, como o percentual da duração do segmento em relação à sua sílaba (*segdurperc*<sup>19</sup>) e a razão da duração da sílaba tônica ou do segmentoônico em relação ao seu par átono (*relative\_increase*), que será detalhada na seção 3.2.4.1. Os dados também foram filtrados para manter apenas a última sílaba de cada palavra, exceto nos casos em que se analisava a coda medial, que estava na penúltima sílaba (ver análise da coda medial na Subseção 3.2.5). Inconsistências, e.g. duplicações de segmentos

<sup>19</sup> Ainda que tivesse sido julgada como necessária no início da elaboração do presente trabalho, esta variável, por fim, não foi utilizada para nenhuma análise.

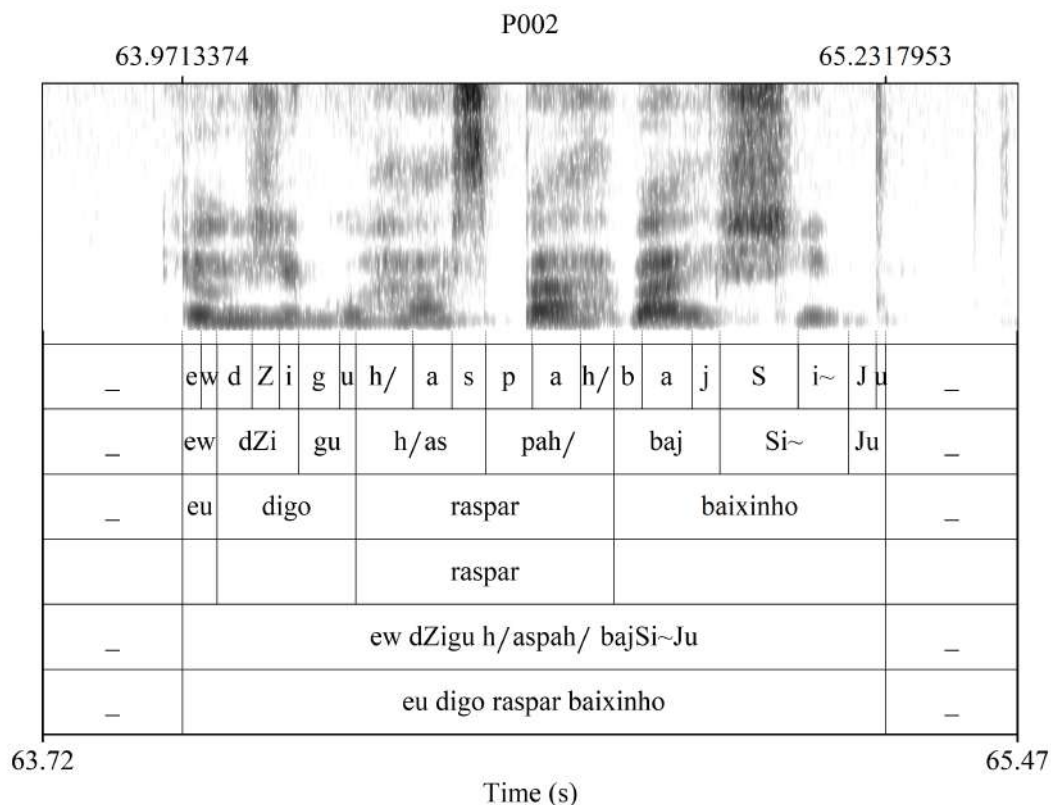


Figura 27 – Panorama geral do TextGrid pronto para a extração de dados.

por sílaba, foram corrigidas ao longo do processo. Essa medida teve como finalidade excluir linhas que não seriam utilizadas nas futuras análises, diminuindo o custo de processamento do computador e, com isso, o tempo de execução do *script*.

### 3.2.4.1 Variável de razão de duração

A variável de razão de duração (*relative\_increase*) foi criada para quantificar a extensão do alongamento da duração das sílabas tônicas em relação às suas contrapartes átonas. As três palavras de cada contexto fonológico, e.g. ‘luva’, ‘uva’ e ‘nova’ vs. ‘madruvá’, ‘Jeová’ e ‘saravá’. Para cada par de sílabas idênticas em estrutura segmental, mas diferentes em acento (*tônica* ou *átona*), foi calculada a razão entre a duração média da sílaba tônica ( $D_{tônica}$ ) e da átona ( $D_{átona}$ ), conforme a fórmula a seguir:

$$\text{razão de duração} = \frac{D_{tônica}}{D_{átona}} \times 100$$

Os valores resultantes foram então integrados a uma nova planilha de dados, mantendo apenas os registros das sílabas tônicas, conforme a Tabela 13.

A Figura 28 compara a apresentação de dados de duração e de razão de alongamento em *boxplots*, a fim de exemplificar a visualização da nova medida e sua interpretação. Pela natureza do seu cálculo, a nova medida possui somente valores acima de zero, i.e. positivos. Valores

Tabela 12 – Descrição dos conjuntos de dados utilizados na análise.

Conjunto de Dados	Características	Medida Principal
Completo	$n$ observações (tônicos e átonos)	Duração bruta em segundos
Reduzido	$n/2$ observações (apenas tônicos)	Razão de duração (valor relativo)

Tabela 13 – Descrição dos conjuntos de dados utilizados na análise.

acima de um indicam que houve um alongamento da duração da sílaba/do segmentoônico em relação à/ao átono; enquanto medidas menores que um apontam que houve um encurtamento. Além disso, essa medida pode ser interpretada como uma porcentagem: por exemplo, 1 = 100%, 1,35 = 135% da duração da átona (ou um aumento de 35%, que é a forma como os resultados serão apresentados).

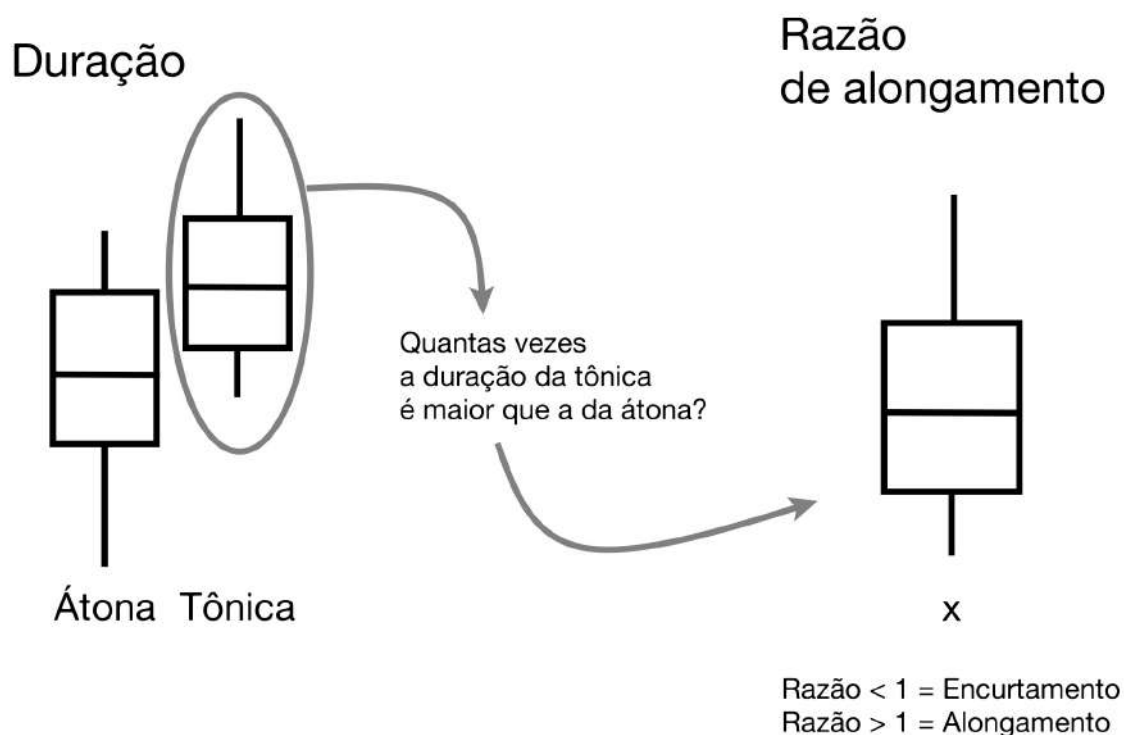


Figura 28 – Ilustração do significado da razão de duração.

O cálculo da nova medida apresenta vantagens, tanto do ponto de vista estatístico quanto interpretativo. Primeiramente, é eliminada a necessidade de incluir a variável referente à tonicidade da sílaba/do segmento diretamente nos modelos estatísticos, pois essa informação já está implícita na construção da variável. Isso simplifica os modelos, o que reduz a sua complexidade e permite um foco maior nas outras variáveis de interesse (como o  $x$  na Figura 28). Além disso, a medida proporciona uma comparação direta entre as durações relativas de tônicas

Estudo	Citação
Abramson e Reo (1990)	“[...] long vowels were on average <b>1.9 times longer</b> than the short vowels.” (p. 83)
Hirata (2004)	“The main acoustic correlate of the short versus long vowels is duration ( <b>ratio 1:2.4–3.2</b> ), although small differences in formants have been reported” (p. 130)
Kang, Yoon e Han (2015)	“[...] long vowels are defined as those produced <b>1.5–2 times longer</b> than short vowels in an identical phonological context.” (p. 480)

Tabela 14 – Medidas de razão (*ratio*) de duração em diferentes estudos. Grifos nossos.

e átonas, destacando claramente o grau de alongamento associado ao acento de palavra. Essa abordagem facilita a análise e a descrição dos resultados, como em *boxplots*, onde as diferenças entre as condições tônica-átona são apresentadas de maneira clara e intuitiva. A ausência de normalização exigiria a comparação de distribuições de durações absolutas, o que dificultaria a visualização e a interpretação dos padrões de interesse.

Deve-se notar que a medida já é utilizada em estudos de fonética e fonologia, inclusive para avaliar o efeito do acento de palavra. Em um estudo bem recente, [Sahkai e Mihkla \(2024\)](#), avaliando as manifestações acústicas do acento em estoniano, apresentam seus resultados para duração como: “*The primary-stressed CV-syllables were on average 9.3 ms (6.5%) longer than the unstressed syllables*” (p. 18). Um outro trabalho exemplo disso é [Paschen, Fuchs e Seifart \(2022\)](#), que, dedicando-se à avaliação do alongamento final em vogais, utiliza a medida de razão (*ratio*) da duração para apresentar seus resultados: “*VV in Dolgan were on average 1.95 times longer than V*” (grifos nossos). Mais especificamente, a medida é apresentada nas tabelas como:

*Table summarizing statistics for all languages without a length contrast. Columns show Language [...], Position (Intercept is Non-final), Number of samples (N), Estimate (b), t-value, p-value, Significance level [...], and Ratio (degree of lengthening between different positions) calculated as mean (Penultimate or Final)/mean (Non-final) for Position and mean (VV)/mean (V) for Vowel Length (ratios are based on non-log values). (PASCHEN; FUCHS; SEIFART, 2022, p. 9, grifos nossos)*

O cálculo da razão, no caso do presente trabalho, como foi apresentado anteriormente, é idêntico ao do trabalho citado, o que indica que outros estudos viram relevância nessa medida normalizada. O uso de “*x times longer*” (“x vezes mais longo(a)”) também será usado neste trabalho, mais especificamente na seção de resultados (Seção 4), como maneira de comunicar o alongamento da unidade acentuada. Além disso, outros trabalhos também se utilizam dessa medida para reportar padrões duracionais em diferentes contextos, como se vê na Tabela 14.

### 3.2.5 Tratamento dos dados de coda medial

Ao concluir o piloto do experimento deste trabalho e analisar os dados de maneira provisória, as medições de duração encontradas para o /s/ em coda fizeram com que se suspeitasse de um efeito de final ou fronteira de palavra na duração desse segmento. Fronteiras de palavra constantemente impactam a duração de segmentos em diversas línguas, sendo o efeito mais comum o alongamento (STERIADE, 1994; BARNES, 2002)<sup>20</sup>, que afetaria grande e diretamente as medições deste trabalho. Essa suspeita levantada faz sentido e poderia ter surgido ainda antes da finalização do piloto, pois já se sabia que somente a última sílaba das palavras do conjunto de dados foram medidas, tendo o desenho do experimento se apoiado na alternância entre palavras paroxítonas (última sílaba átona) e oxítonas (última sílaba tônica) para avaliar o efeito do acento de palavra.

Para que isso fosse avaliado, então, 18 novas palavras foram adicionadas ao conjunto de dados, conforme a Tabela 15.

<b>Coda medial tônica</b>	<b>Coda medial átona</b>
“raspa”	“raspar”
“caspá”	“caspar”
“aspa”	“aspar”
“basta”	“bastar”
“pasta”	“gastar”
“vasta”	“pastar”
“casca”	“bastar”
“lasca”	“lascar”
“masca”	“mascar”

Tabela 15 – Palavras adicionadas ao conjunto de dados para avaliação do efeito de coda medial.

O /s/ em coda medial passou pelo mesmo processamento do conjunto de dados original, e depois suas medições foram adicionadas ao restante dos dados. É importante apontar que somente o segmento em coda foi medido e nenhuma outra parte das novas palavras. Essa decisão foi tomada para evitar desequilíbrios no desenho original do experimento, que poderia ser comprometido com a inclusão de novas consoantes em ataque simples, em posição de início/fronteira de palavra. A distribuição de [s] nos dados pode ser visualizada na Tabela 16.

Tabela 16 – Distribuição das ocorrências de [s] em posições mediais e finais de palavras, considerando a tonicidade.

<b>Tonicidade</b>	<b>Em posição medial</b>	<b>Em posição final</b>
Não acentuada	100	177
Acentuada	104	272

Propriedade	Descrição	Valor
Total de linhas	Número total de observações	7860
Total de colunas	Número total de variáveis	22
Tipos de variáveis	factor (17), numeric (5)	-
Variáveis categóricas	Exemplo: part, segtype, stress	-
Variáveis numéricas	Exemplo: segdur, sylldur, worddur	-

Tabela 17 – Resumo geral dos dados.

Variável	Média	Mediana	Desvio padrão	Mínimo	Máximo	N únicos
segdur	0.0902	0.0780	0.0464	0.0157	0.270	-
sylldur	0.266	0.258	0.0877	0.0600	0.496	-
worddur	0.507	0.500	0.103	0.204	0.799	-
segtype	-	-	-	-	-	2 (C, V)

Tabela 18 – Estatísticas resumidas das variáveis numéricas e categóricas.

### 3.2.6 Resumo descritivo

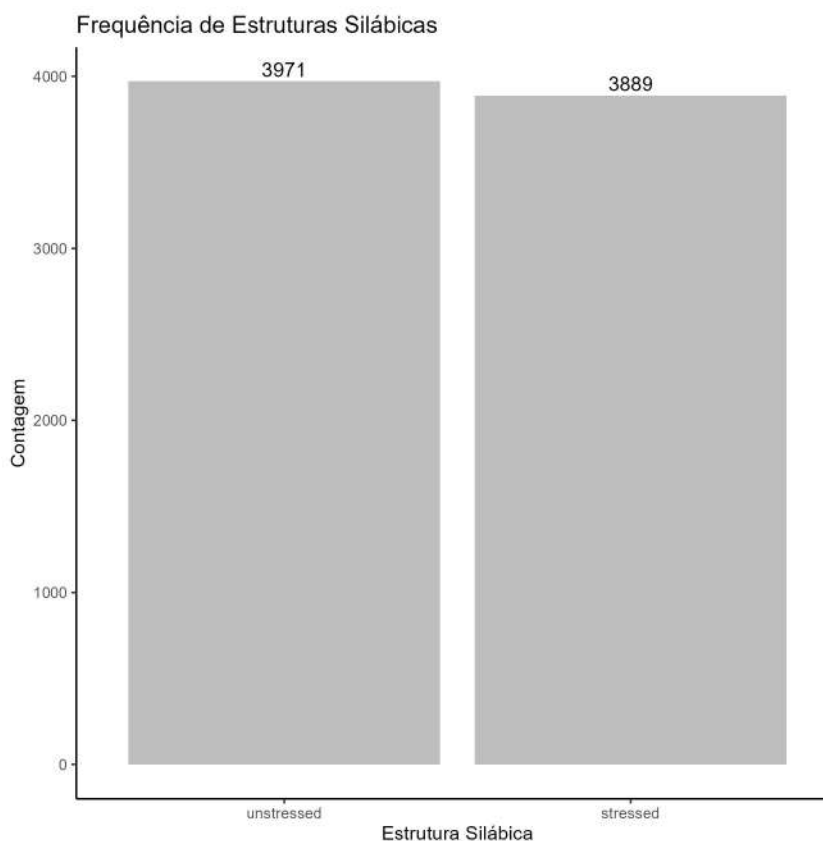


Figura 29 – Distribuição da tonicidade dos segmentos.

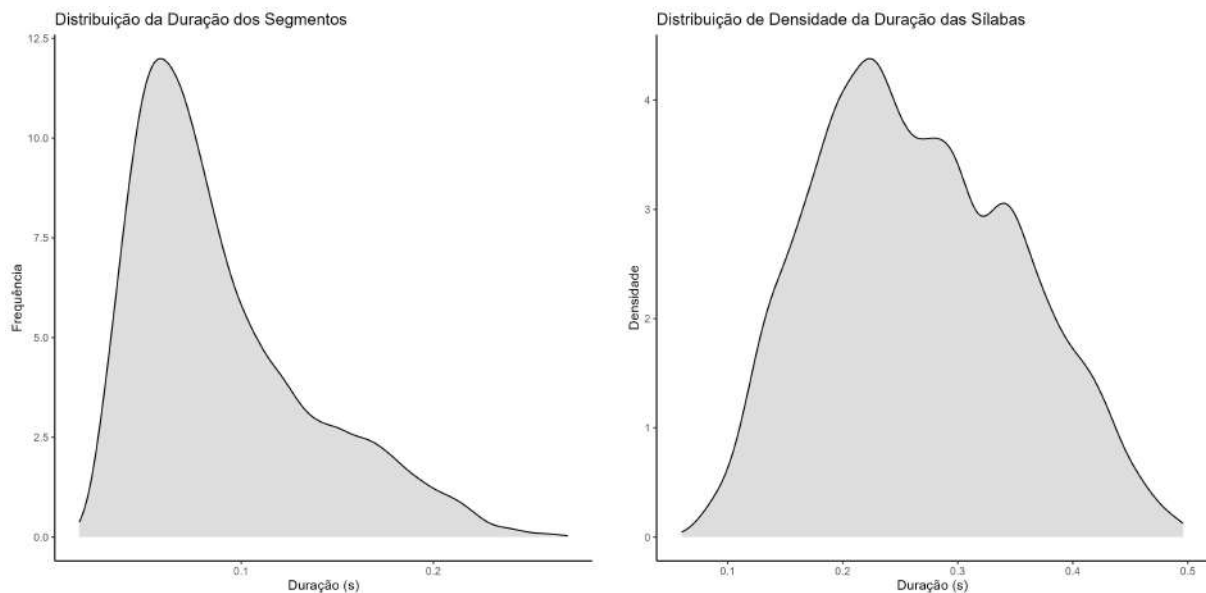
A análise inicial dos dados revela as seguintes características principais: os segmentos apresentam uma média de duração de 0,09 segundos (30a), enquanto as sílabas duram em média

<sup>20</sup> Ver [Rouch e Lunden \(2020\)](#) para uma revisão.

Tabela 19 – Descrição das Variáveis.

Tipo	Variável	Descrição e Valores
Qualitativas	syllstruc	Molde silábico (CV, CCV, CVC e CCVC).
	sylltree	Constituinte do segmento na sílaba (ons, ons1, ons2, N e cd).
	syllpos	Posição da sílaba na palavra.
	segtype	Tipo de segmento.
	wordstresstype	Tipo de acento.
	artp	Ponto de articulação (glotal, alveolar, etc.).
	artm	Modo de articulação (fricativa, oclusiva, nasal).
Quantitativas	segdur	Duração do segmento (média: 0,09s, variação: 0,02s–0,27s).
	relative_increase	Relação entre duração do tônico e átono no mesmo contexto.
	sylldur	Duração da sílaba (média: 0,26s, variação: 0,06s–0,56s).
	segdurperc	Percentual da duração do segmento na sílaba (média: 35,20%).
	wordsylln	Número de sílabas por palavra (2 a 4).
	syllsegn	Número de segmentos por sílaba (2 a 4).
	segsyllpos	Posição do segmento na sílaba (1 a 4).

0,25 segundos (30b) e as palavras, 0,47 segundos. Além disso, a maioria das variáveis categóricas apresenta uma distribuição equilibrada, com exceção de algumas, como *wordsylln* e *syllsegn*, que têm valores centrados em 2 e 3, respectivamente. Essas análises iniciais serviram como base para a subsequente aplicação de modelos estatísticos.



(a) Distribuição da duração de segmentos.

(b) Distribuição da duração de sílabas.

Figura 30 – Distribuições das durações de segmentos e sílabas no conjunto de dados.

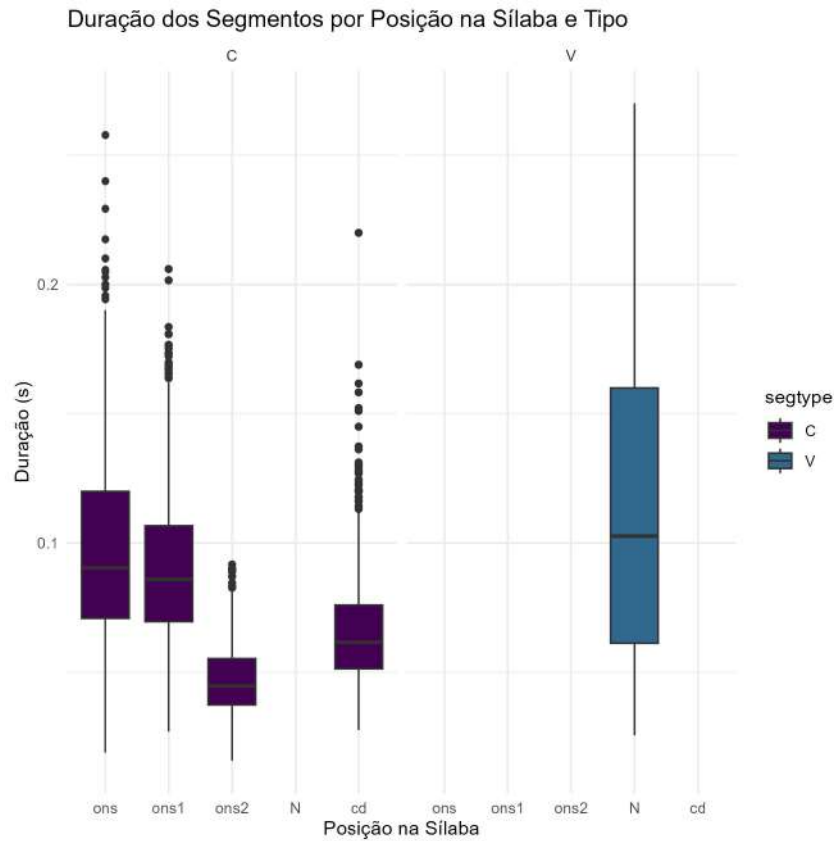
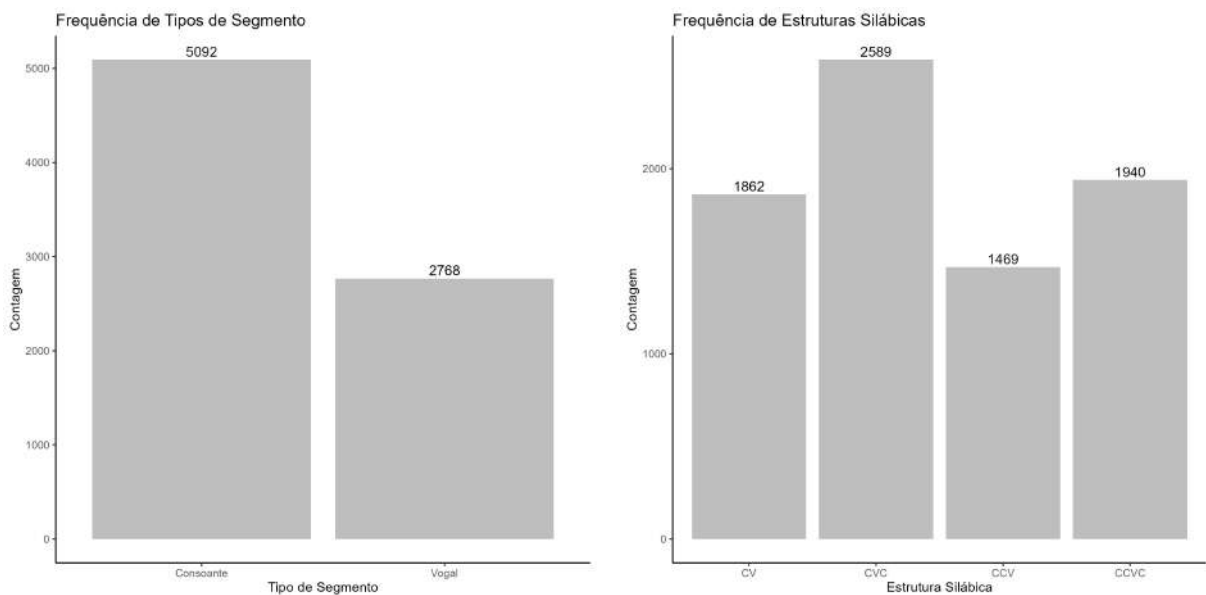


Figura 31 – Distribuição da duração de consoantes e vogais no conjunto de dados.



(a) Contagem de consoantes e vogais.

(b) Contagem de sílabas de cada molde.

Figura 32 – Frequências de tipos de segmentos e moldes silábicos no conjunto de dados.

### 3.2.7 Avaliação estatística

A avaliação estatística dos dados foi conduzida utilizando modelos mistos lineares no ambiente de programação R, com os pacotes `lme4`, `MuMIn`, `performance` e `multcomp`. A

variável dependente foi a razão de duração<sup>21</sup>, enquanto as variáveis independentes variaram para cada caso (sílabas, consoante em primeira posição do ataque, tepe, vogal /a/ e /s/ em coda). Um efeito aleatório foi incluído para o fator “participante”, permitindo modelar a variabilidade entre os indivíduos.

Para selecionar o modelo de melhor ajuste, foi empregada a função `dredge` do pacote `MuMIn`, que realiza, a partir do maior modelo (global), uma busca exaustiva por todos os modelos possíveis derivados das combinações dos preditores especificados no modelo inicial.

O Critério de Informação de Akaike (AIC) (AKAIKE, 1974) foi utilizado como métrica para selecionar o modelo que melhor se ajusta aos dados, balanceando a qualidade do ajuste e a complexidade do modelo. Caso a razão  $n/k$  (número de observações dividido pelo número de parâmetros no modelo) fosse menor que 40, o critério passaria a ser o AIC corrigido (AICc), mais ideal para um número pequeno de dados (HURVICH; TSAI, 1989). No entanto, em nenhum ajuste de modelo foi necessário recorrer à medida ajustada, devido ao alto volume de dados por variável preditora.

A seleção foi baseada no Critério de Informação de Akaike (AIC)<sup>22</sup>, que penaliza modelos mais complexos para evitar sobreajuste. A função `dredge()` gerou uma tabela de seleção de modelos, apresentando os valores de AIC para cada combinação de preditores; aquele modelo com o menor AIC foi escolhido.

Após a seleção do melhor modelo, este foi submetido a verificações de diagnóstico, incluindo a avaliação da normalidade dos resíduos e da homogeneidade das variâncias, utilizando ferramentas do pacote `performance`. Um histograma foi gerado para inspecionar visualmente a normalidade dos resíduos. Por fim, foi realizada uma análise *post hoc* (Teste de Turkey) para medir as diferenças entre os níveis dos efeitos fixos, com a função `glht` do pacote `multcomp`, utilizando o método de correção de Holm para ajustar os valores de significância e do valor-p.

<sup>21</sup> Como já foi tratado na Subseção 3.2.4.1, dessa forma se conseguiu omitir a variável `stress` das variáveis predictoras, diminuindo o custo estatístico do modelo e, principalmente, tornando a sua interpretação menos complexa.

<sup>22</sup> O critério passaria a ser AICc, caso  $n$  (número de observações) dividido por  $k$  (número de variáveis predictoras) fosse menor que 40.

## 4 Resultados

A seguir, são apresentados os principais resultados, seguindo uma lógica parecida para todas as seções: tabelas apresentarão os resultados do modelo linear misto selecionado e do teste de Turkey das variáveis preditoras do modelo; depois serão mostrados dois gráficos de duração bruta em milissegundos (agrupados por molde e por tonicidade) e um referente à razão de duração<sup>1</sup> da sílaba/do segmento. Em alguns casos, quando alguma variável foi estatisticamente significativa, será adicionada alguma figura relevante para a verificação visual da diferença encontrada pelo modelo linear misto.

Um resumo dos valores de alongamento da sílaba tônica, assim como foi proposto que seria feito na Figura 2, está a seguir na Figura 33.

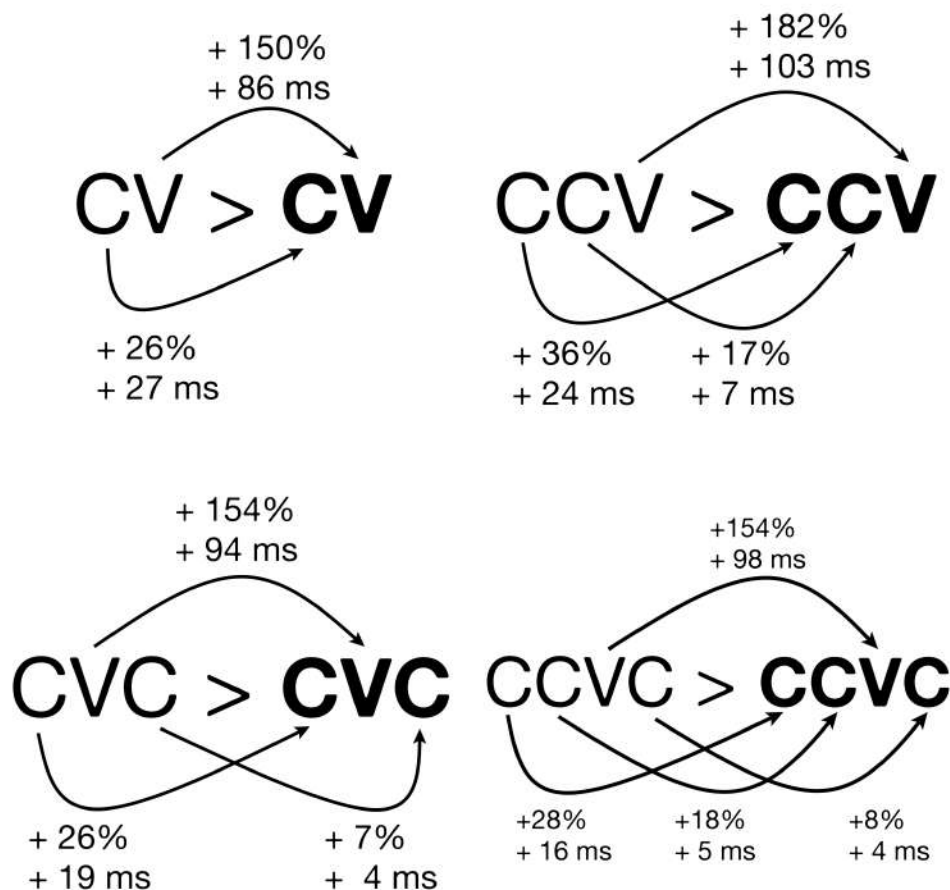


Figura 33 – Resumo dos resultados de alongamento dos segmentos (molde silábico por posição na sílaba). O negrito marca a sílaba tônica.

As interpretações e discussões causadas pelos seguintes resultados serão apresentados e desenvolvidos na próxima seção do texto (Seção 5).

<sup>1</sup> Ver mais informações sobre o cálculo dessa medida na Subseção 3.2.4.1

## 4.1 Duração da sílaba

A duração da sílaba tônica foi maior em todas as estruturas silábicas (Figuras 34 e 35a). A razão de alongamento das sílabas tônicas variou significativamente entre os diferentes moldes, e esse efeito foi estatisticamente significativo, como será apresentado a frente.

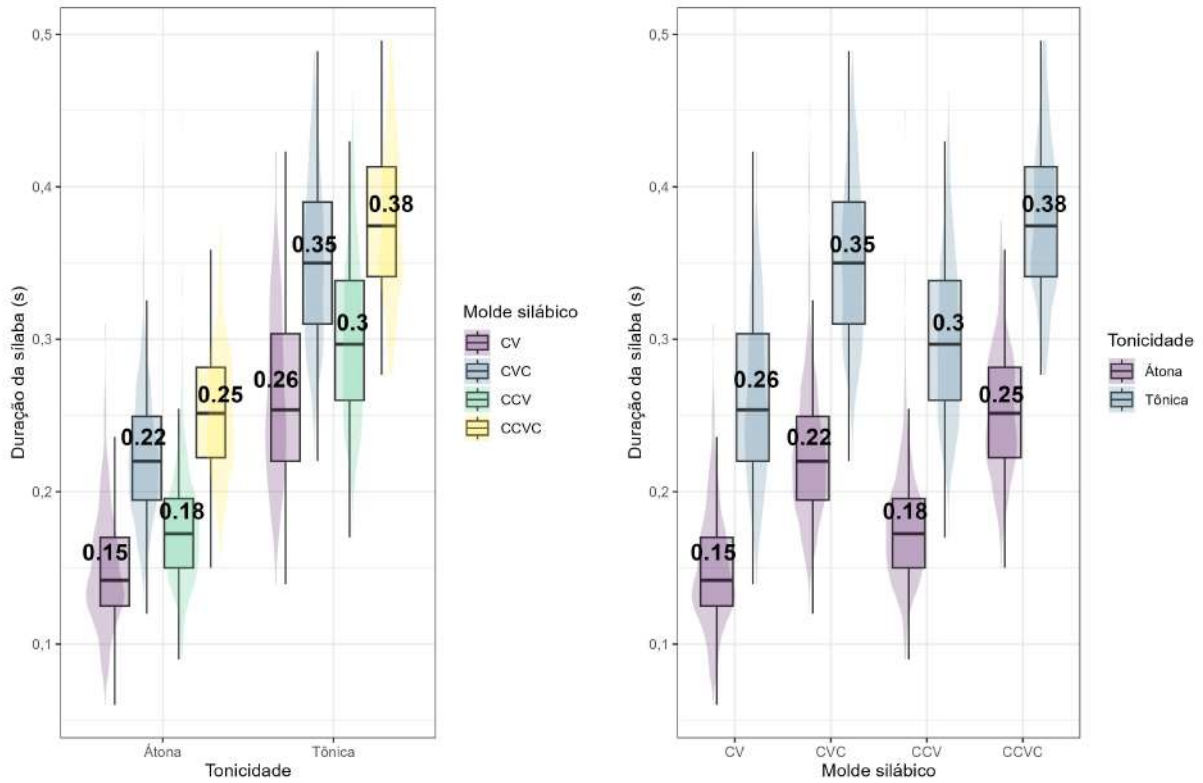
Quanto ao ajuste do modelo linear misto, o melhor selecionado com base em AIC foi o `lmer(relative_increase ~ syllstruc + voz + (1 | part), data = data_syll)`, que está descrito na Tabela 20. Em resumo, os valores previstos pelo modelo apontam que sílabas CV tônicas medem 1,8 vezes mais que as átonas, nos outros moldes, também há esse alongamento, porém em menor grau, como se vê a seguir: sílabas CVC tônicas medem 1,6 vezes mais (\*\*); CCV, 1,7 (\*); e CCVC, 1,53 (\*\*\*) (Figura 35a). O vozeamento da consoante no ataque também influenciou a razão de alongamento (Figura 35b), e, de acordo com o modelo, há um efeito ligeiramente menos significativo: sílabas iniciadas por consoantes desvozeadas medem 1,61 vezes mais (\*\*), e aquelas iniciadas por consoantes vozeadas, 1,71 vezes mais (\*).

Tabela 20 – Resultados do modelo misto para razão de duração das sílabas.

Efeito Fixo	Estimativa	Erro Padrão	Valor-t	Significância
Intercepto	1,75359	0,03921	44,721	***
syllstrucCVC	-0,20883	0,01618	-12,907	***
syllstrucCCV	-0,09776	0,01897	-5,154	***
syllstrucCCVC	-0,27061	0,01901	-14,237	***
vozvoiced	0,08686	0,01306	6,649	***
<i>Efeitos Aleatórios</i>				
Intercepto (participantes): Variância = 0,146, Desvio Padrão = 0,382				
Resíduo: Variância = 0,319, Desvio Padrão = 0,565				
<b>Informações do Modelo</b>	Número de observações: 1346			
	Número de grupos (participantes): 14			

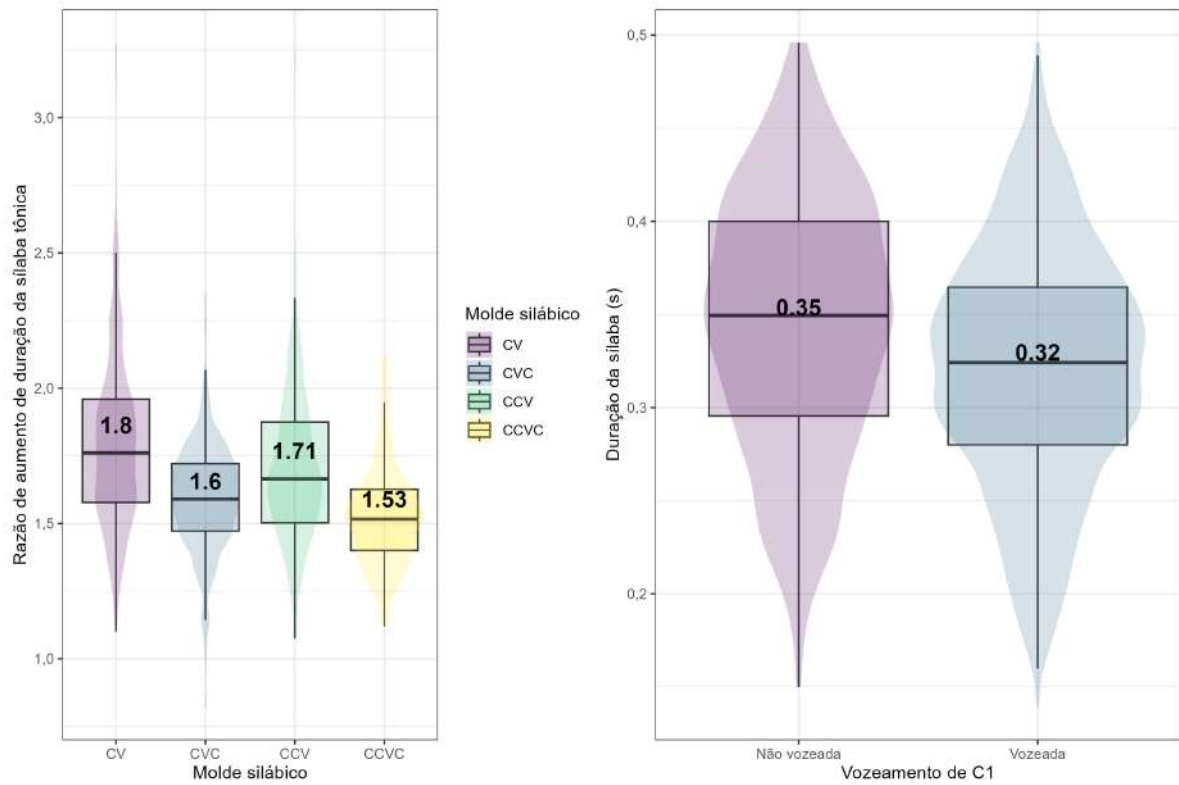
Tabela 21 – Resultados dos Contrastes de Tukey (Método de Holm) para o alongamento da sílaba tônica.

Comparação	Estimativa	Erro Padrão	Valor z	p-ajustado	Significância
CVC - CV	-0,20883	0,01618	-12,907	$< 2 \times 10^{-16}$	***
CCV - CV	-0,09776	0,01897	-5,154	$5,11 \times 10^{-7}$	***
CCVC - CV	-0,27061	0,01901	-14,237	$< 2 \times 10^{-16}$	***
CCV - CVC	0,11107	0,01940	5,726	$3,09 \times 10^{-8}$	***
CCVC - CVC	-0,06178	0,01943	-3,180	0,00147	**
CCVC - CCV	-0,17285	0,02182	-7,922	$9,77 \times 10^{-15}$	***
Voiced - Unvoiced	0,08686	0,01306	6,649	$2,95 \times 10^{-11}$	***



(a) Duração da sílaba agrupada por tonicidade. (b) Duração da sílaba agrupada por molde silábico.

Figura 34 – Comparação da duração da sílaba em diferentes contextos.



(a) Razão de duração da sílaba por molde silábico. (b) Razão de duração da sílaba por vozeamento da consoante em ataque.

## 4.2 Duração dos segmentos

### 4.2.1 O núcleo da sílaba: /a/

O alongamento das vogais nas sílabas tônicas variou pouco entre os diferentes moldes (cf. Figura 36), mas esse efeito foi estatisticamente significativo para alguns tipos de moldes. A Tabela 22 mostra os resultados do melhor modelo, i.e. aquele ajustado somente com o molde silábico como variável preditora. Pôde-se perceber que somente o molde CCV foi relevante em comparação ao intercepto do ajuste (CV).

Tabela 22 – Resultados do modelo misto para razão de duração das vogais.

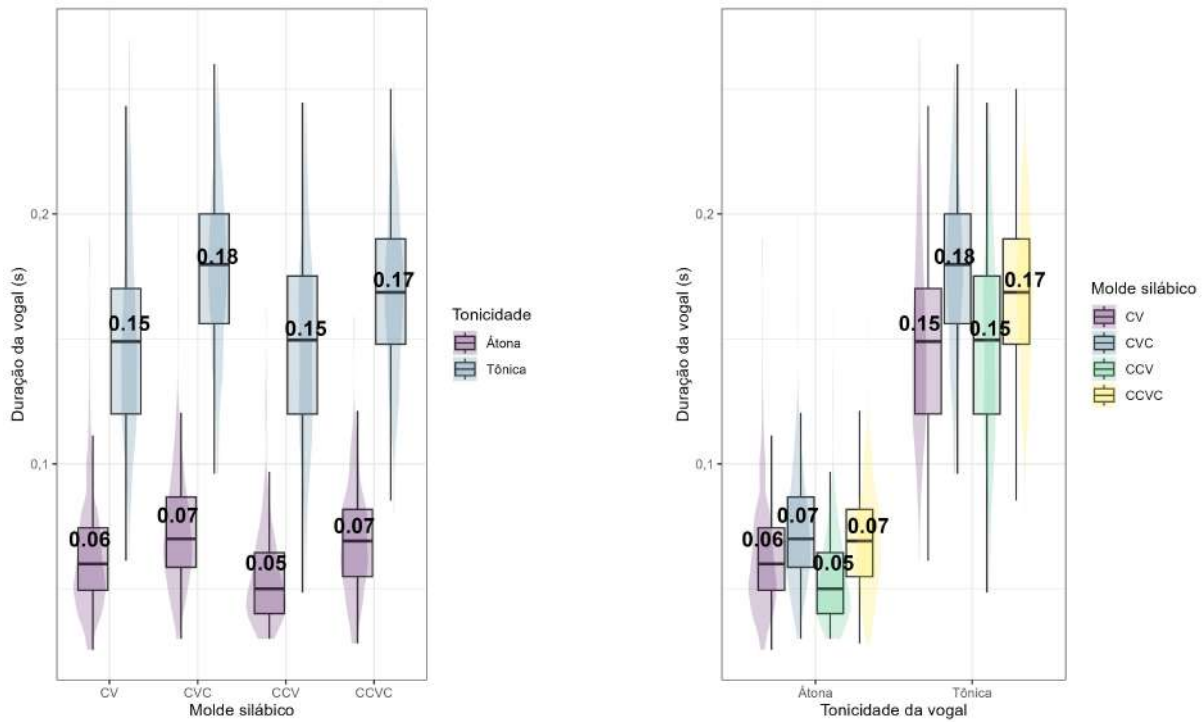
Efeito Fixo	Estimativa	Erro Padrão	Valor-t	Significância
Intercepto	2,502	0,1049	23,850	***
syllstrucCVC	0,050	0,04193	1,209	
syllstrucCCV	0,318	0,04919	6,468	***
syllstrucCCVC	0,047	0,04929	0,966	
<i>Efeitos Aleatórios</i>				
Intercepto (participantes): Variância = 0,1423, Desvio Padrão = 0,3773				
Resíduo: Variância = 0,3792, Desvio Padrão = 0,6158				
<b>Informações do Modelo</b>	Número de observações: 1346			
	Número de grupos (participantes): 14			

Partindo para a avaliação da comparação entre grupos, neste caso somente entre os moldes silábicos, apenas algumas diferenças entre grupos são significantes. O teste de Tukey (Tabela 23) aponta que o núcleo do molde CCV se alonga 0,31 vezes (i.e. 31%) mais do que aquele em molde CV e 26% mais que o em CVC; já o núcleo de uma sílaba CCV se alonga 27% mais do que o aquele em uma CCVC. De maneira geral, aquele núcleo que mais se alonga é aquele presente em uma sílaba de ataque complexo e sem coda, i.e. CCV, conforme pode ser visualizado na Figura 37.

Tabela 23 – Resultados do Teste de Comparações Múltiplas (Tukey) com Ajuste de Holm.

Comparação	Estimativa	Erro Padrão	Valor p
CVC - CV	0,051	0,042	0,680
CCV - CV	0,318	0,049	< 0,001***
CCVC - CV	0,048	0,049	0,680
CCV - CVC	0,267	0,050	< 0,001***
CCVC - CVC	-0,003	0,050	0,951
CCVC - CCV	-0,271	0,057	< 0,001***

0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 ',' 0,1 ' ' 1



(a) Duração de /a/ agrupada por molde silábico.

(b) Duração de /a/ agrupada por tonicidade.

Figura 36 – Comparação da duração de /a/ em diferentes contextos.

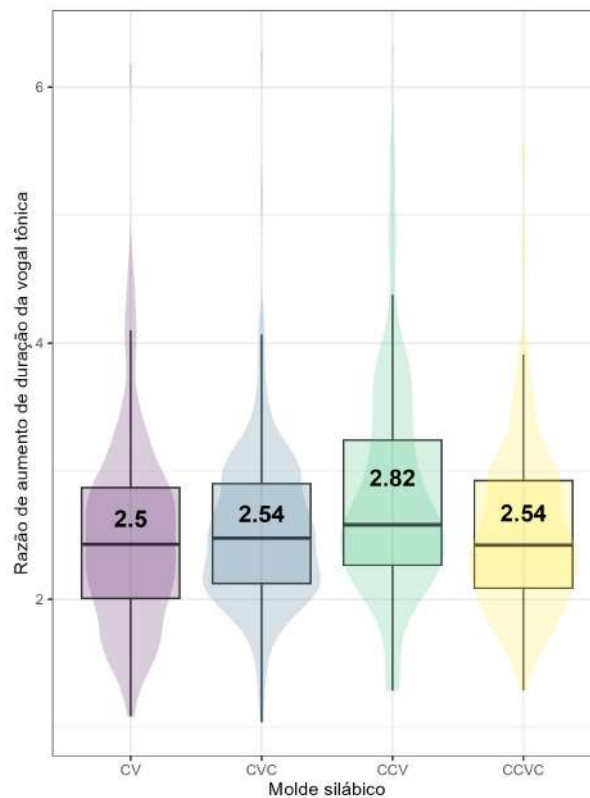


Figura 37 – Razão de duração /a/ por molde silábico.

## 4.2.2 As partes não nucleares da sílaba

### 4.2.2.1 As consoantes em ataque simples ou ramificado

O alongamento das consoantes em ataque (simples ou ramificado) nas sílabas tônicas esteve presente em todos os moldes, porém em diferentes intensidades. O melhor modelo selecionado com base em AIC foi o `lmer(relative_increase ~ syllstruc + voz + (1 | part), data = data_seg_c)`.

É importante citar que as consoantes nasais ([m] e [n]) foram excluídas do conjunto de dados utilizado para rodar o modelo linear misto desta parte da análise devido a um problema de multicolinearidade com outras variáveis do modelo. As vogais nasais estavam restritas a estruturas silábicas com ataque simples (CV e CVC), o que gerava uma correlação forte com a variável que codifica a estrutura silábica. Além disso, as consoantes nasais também afetavam a variável que define a posição do consoante dentro da sílaba, uma vez que essas consoantes sempre apareciam no mesmo constituinte silábico, em ataque simples, codificado `ons`. Esse tipo de multicolinearidade é problemático para modelos lineares mistos, pois a redundância entre variáveis pode inflar as variâncias dos coeficientes estimados, dificultando a interpretação e a precisão dos resultados. A exclusão das consoantes nasais ajudou a reduzir essa correlação e a melhorar a estabilidade do modelo, permitindo uma análise mais confiável das relações entre as variáveis. De toda forma, essas consoantes ainda foram mantidas nos gráficos que se seguem, na esperança de sua visualização ser útil.

A correlação direta entre as variáveis de estrutura silábica e constituinte silábico fez com que o modelo não fosse ajustado as considerando. Isso acontece, pois as estruturas CV e CVC já implicam que a sua consoante em ataque está sozinha, i.e. está posicionada em um ataque simples (`ons`). De forma complementar, sílabas CCV e CCVC estão diretamente relacionadas a consoantes marcadas como `ons1`, i.e. encabeçam um ataque ramificado.

Como se vê nas Tabelas 24 e 25, o efeito do molde silábico no alongamento da consoante que inicia o ataque (i.e. C1) foi significativo para os moldes CV, CVC e CCVC e não significativo para o CCV. De modo geral, segundo os valores estimados do modelo, sílabas CV apresentaram um aumento de 36,2% (\*\*); CVC: 17,3% (\*\*); CCV: 26,7% (ns); e CCVC: 18,3% (\*).

Quanto ao vozeamento, aquelas consoantes não vozeadas se alongaram 24,1%, e as vozeadas, 23%. O modo de articulação também teve efeitos significantes no alongamento: fricativas tônicas são 39,6% (\*\*\*) maiores (como ilustra a Figura 39); nasais: 26,1% (\*\*\*) (Figura 41); e oclusivas: 27,4% (\*\*\*) (Figura 40).

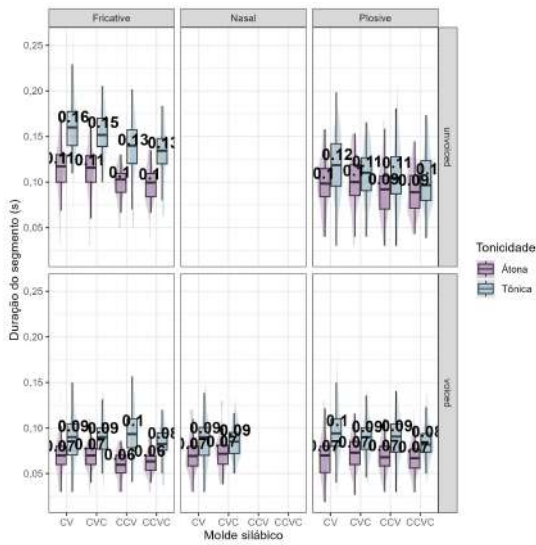
Tabela 24 – Resultados do modelo misto para razão de duração das consoantes em primeira posição de ataque.

<b>Efeito Fixo</b>	<b>Estimativa</b>	<b>Erro Padrão</b>	<b>Valor-t</b>	<b>Significância</b>
Intercepto	1,38976	0,03908	35,559	***
artmNasal	-0,16896	0,03195	-5,289	***
artmPlosive	-0,12726	0,01821	-6,990	***
syllstrucCVC	-0,09855	0,02059	-4,786	***
syllstrucCCV	0,01294	0,02486	0,520	
syllstrucCCVC	-0,05515	0,02503	-2,204	*
vozvoiced	0,08868	0,01734	5,113	***
<i>Efeitos Aleatórios</i>				
Intercepto (participantes): Variância = 0,01623, Desvio Padrão = 0,1274				
Resíduo: Variância = 0,09134, Desvio Padrão = 0,3022				
<b>Informações do Modelo</b>	Número de observações: 1346			
	Número de grupos (participantes): 14			

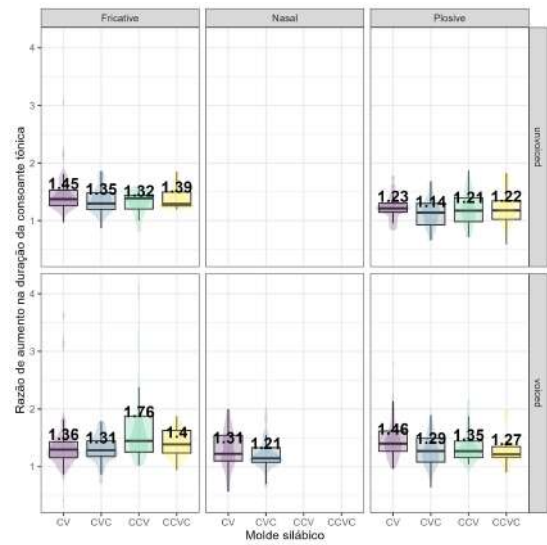
Tabela 25 – Resultados dos Contrastes de Tukey (Método de Holm).

<b>Comparação</b>	<b>Estimativa</b>	<b>Erro Padrão</b>	<b>Valor z</b>	<b>Valor-p ajust.</b>
<b>Modo de Articulação</b>				
Nasal - Fricativa	-0,16896	0,03195	-5,289	$2,47 \times 10^{-7}$ ***
Plosiva - Fricativa	-0,12726	0,01821	-6,990	$8,22 \times 10^{-12}$ ***
Plosiva - Nasal	0,04170	0,03193	1,306	0,192
<b>Molde Silábico</b>				
CVC - CV	-0,09855	0,02059	-4,786	$1,02 \times 10^{-5}$ ***
CCV - CV	0,01294	0,02486	0,520	0,6028
CCVC - CV	-0,05515	0,02503	-2,204	0,0826 .
CCV - CVC	0,11149	0,02537	4,394	$5,56 \times 10^{-5}$ ***
CCVC - CVC	0,04339	0,02551	1,701	0,1779
CCVC - CCV	-0,06809	0,02778	-2,451	0,0569 .
<b>Vozeamento</b>				
vozeada - não vozeada	0,08868	0,01734	5,113	$3,17 \times 10^{-7}$ ***

0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 ',' 0,1 ' ' 1

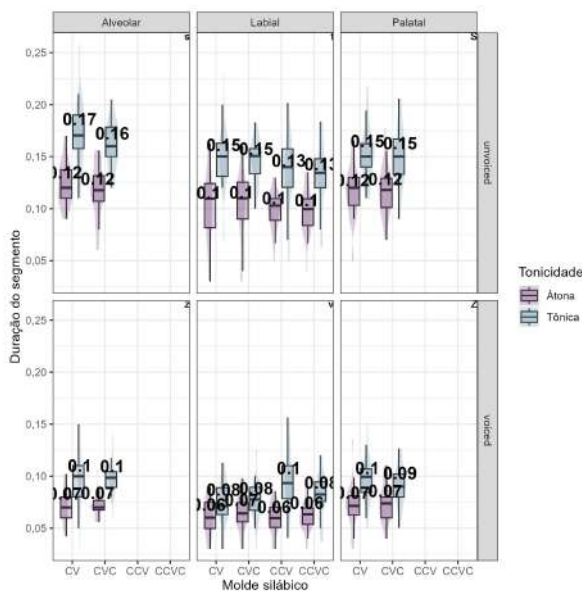


(a) Duração de consoantes separada por modo de articulação e vozeamento.

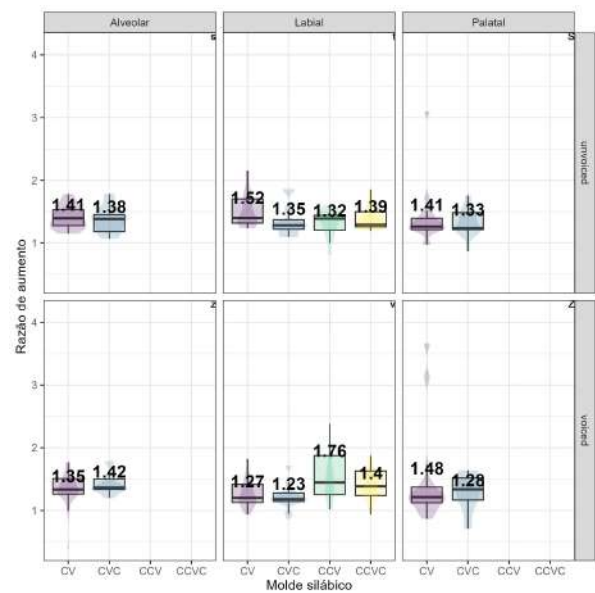


(b) Razão de duração de consoantes separada por modo de articulação e vozeamento.

Figura 38 – Efeito do acento de palavra medido em duração bruta e razão de duração, com separação por modo de articulação e vozeamento.

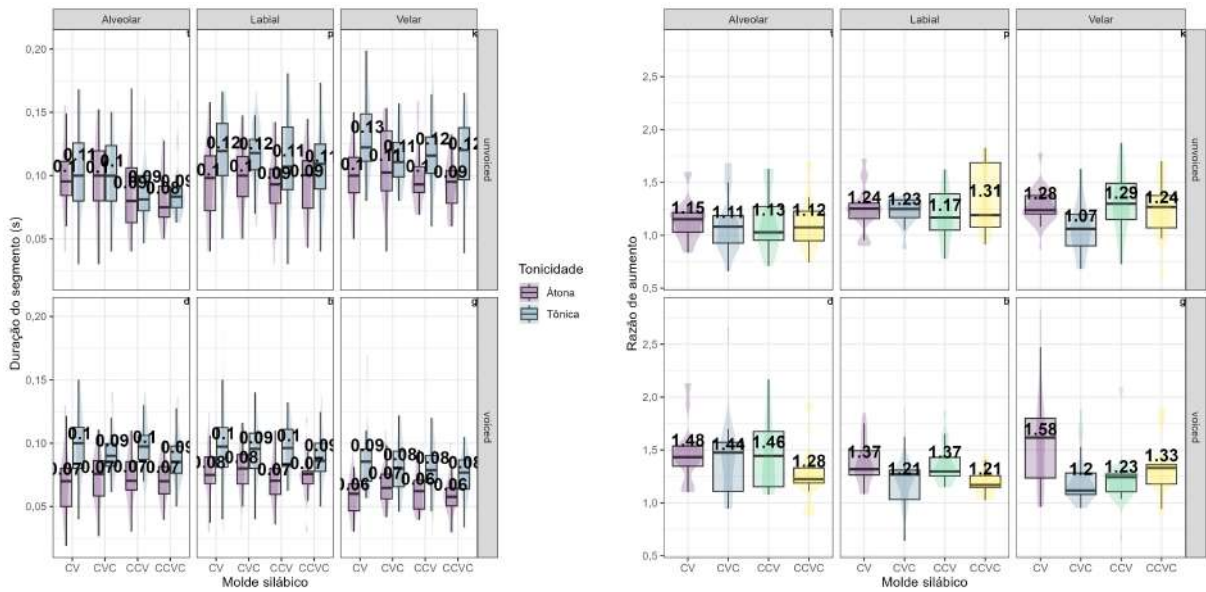


(a) Duração das fricativas.



(b) Razão de duração das fricativas.

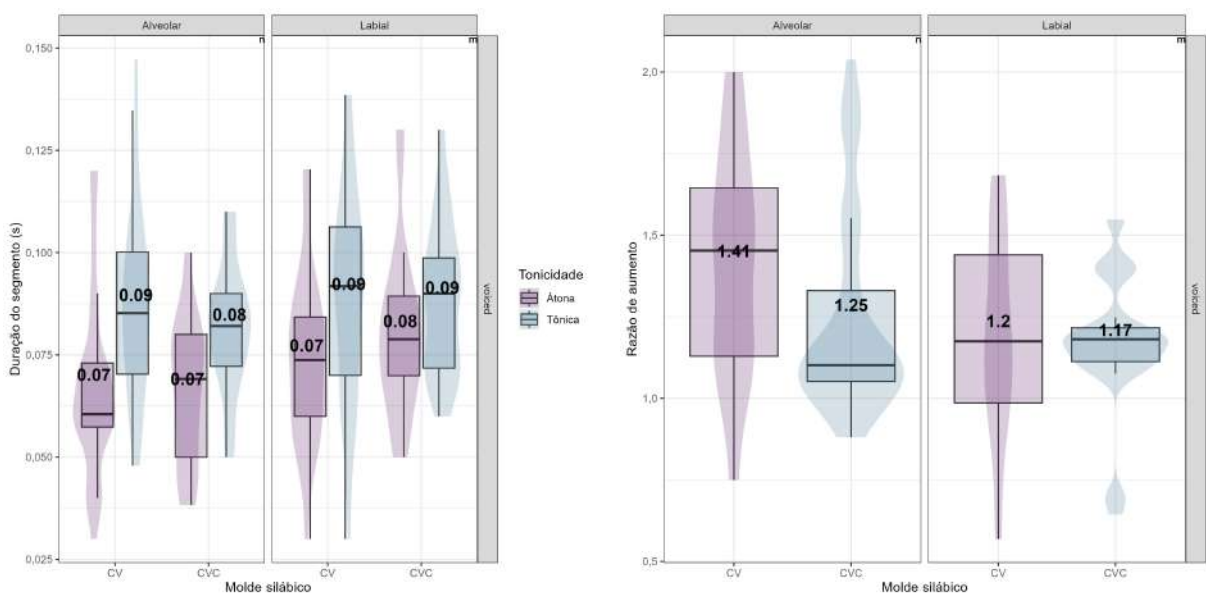
Figura 39 – Gráficos de análise de fricativas.



(a) Duração das oclusivas.

(b) Razão de duração das oclusivas.

Figura 40 – Gráficos de análise de oclusivas.



(a) Duração das nasais.

(b) Razão de duração das nasais.

Figura 41 – Gráficos de análise de nasais.

#### 4.2.2.2 A consoante em ramificação de ataque: [r]

É relevante, antes de tudo, afirmar que a duração bruta do tepe foi afetada em algum nível (e de maneira estatisticamente relevante) pelo acento de palavra (Figura 42a), conforme aponta o modelo linear misto  $\text{lmer}(\text{segdur} \sim \text{stress} + (1 | \text{part}), \text{data} = \text{data}_r)$ , que, diferente dos modelos que estavam sendo ajustados até aqui, tem como variável-resposta a duração bruta, e não a razão de alongamento. Conforme será apresentado e discutido, não houve um efeito de alongamento aparente no tepe, o que faz a análise da duração bruta ser mais

proveitosa para análise e discussão. Os resultados do melhor modelo se encontram na Tabela 26. Em média, o tepe em sílaba acentuada durou 5 ms, i.e. 13% de aumento.

Tabela 26 – Resultados do modelo misto para duração do tepe em ramificação de ataque complexo.

Efeito Fixo	Estimativa	Erro Padrão	Valor-t	Significância
Intercepto	0,04395	0,00161	27,259	***
stressed	0,00650	0,00076	8,506	***

*Efeitos Aleatórios*

Intercepto (participantes): Variância = 0,00003225, Desvio Padrão = 0,005679  
 Resíduo: Variância = 0,0001413, Desvio Padrão = 0,011889

Informações do Modelo	Valor
Número de observações:	975
Número de grupos (participantes):	14

Quanto à razão de alongamento do tepe ao longo dos moldes silábicos (Figura 42b) não foi relevante. Assim, o melhor modelo selecionado foi o nulo, pois foi aquele de menor AIC no método usado.

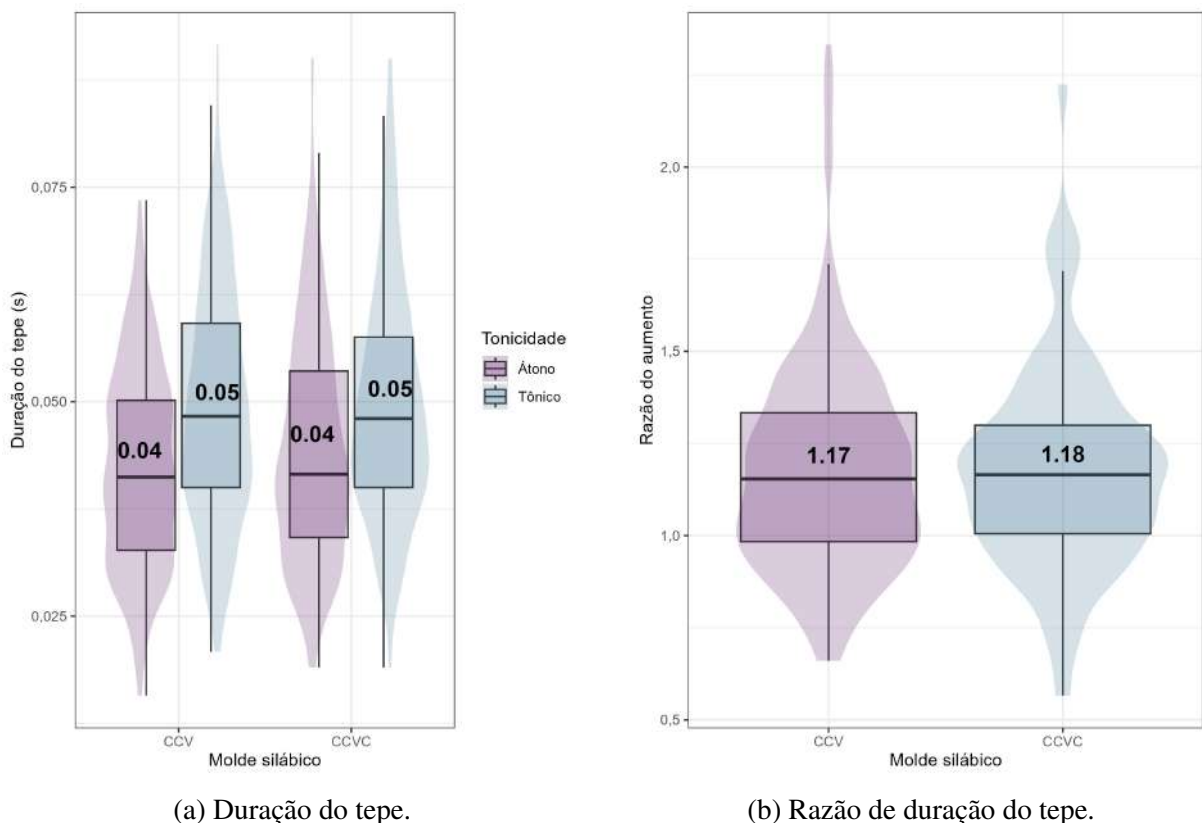


Figura 42 – Comparação entre a duração e a razão de duração do tepe.

#### 4.2.2.3 A consoante em coda: /s/

Devido às diferentes realizações do gesto em coda, foram ajustados dois modelos, considerando se o segmento é [s] ou [z]. O [s] em ataque se alonga 41,1%, e o em coda, 12,8%

(Tabela 27). No caso de [z], em ataque se alonga 36,9%, e em coda, 2,6% (Tabela 28). Em ambos os casos, percebe-se como o ataque é bem mais alongado pelo acento de palavra do que a coda (cf. Figura 43).

Tabela 27 – Resultados do modelo misto para razão de duração de [s].

<b>Efeito Fixo</b>	<b>Estimativa</b>	<b>Erro Padrão</b>	<b>Valor-t</b>	<b>Significância</b>
Intercepto	1,41124	0,04331	32,59	***
sylltreecd	-0,28336	0,04054	-6,99	***
<i>Efeitos Aleatórios</i>				
Intercepto (participantes): Variância = 0,01330, Desvio Padrão = 0,1153				
Resíduo: Variância = 0,05736, Desvio Padrão = 0,2395				
<b>Informações do Modelo</b>	Número de observações: 245			
	Número de grupos (participantes): 14			

Tabela 28 – Resultados do modelo misto para razão de duração de [z].

<b>Efeito Fixo</b>	<b>Estimativa</b>	<b>Erro Padrão</b>	<b>Valor-t</b>	<b>Significância</b>
Intercepto	1,36953	0,03982	34,40	***
sylltreecd	-0,34359	0,03246	-10,59	***
<i>Efeitos Aleatórios</i>				
Intercepto (participantes): Variância = 0,01054, Desvio Padrão = 0,1027				
Resíduo: Variância = 0,04916, Desvio Padrão = 0,2217				
<b>Informações do Modelo</b>	Número de observações: 429			
	Número de grupos (participantes): 14			

### O /s/ em coda medial

Assim como consta na Subseção 3.2.5, foi necessário que 18 novas palavras fossem adicionadas ao conjunto de dados do experimento de leitura, a fim de que fosse avaliado o efeito da fronteira final de palavra na duração. As medições de duração dessa porção dos dados mostram que o [s]<sup>2</sup> tônico em coda medial é menor do que aquele em coda medial átona (Figuras 44 e 45). Logicamente, a média da razão de duração se mantém, portanto, menor que um (cf. Figura 44b), indicando que o denominador (duração do segmento átono) é maior que o numerador (tônico) no cálculo da medida de alongamento<sup>3</sup>.

Os resultados do modelo misto (Tabela 29) mostram que a razão de duração da coda medial é, em média, 0,8746, indicando que o /s/ em coda tônica é mais curto do que em coda átona. O teste de Tukey (Tabela 30) confirma que essa diferença é significativa ( $z = 8,07$ ,  $p < 0,001$ ).

<sup>2</sup> As codas em posição de meio de palavra não passaram pelo mesmo processo de vozeamento que algumas daquelas em posição final, conforme apresentado na Subseção 3.2.3.1. Por esse motivo, os dados que correspondiam ao som vozeado ([z]) foram desconsiderados para a construção dos gráficos que comparam ataque e coda. Dessa forma, consegue-se manter a comparação apenas entre sons não-vozeados.

<sup>3</sup> Ver discussão na Subseção 3.2.4.1.

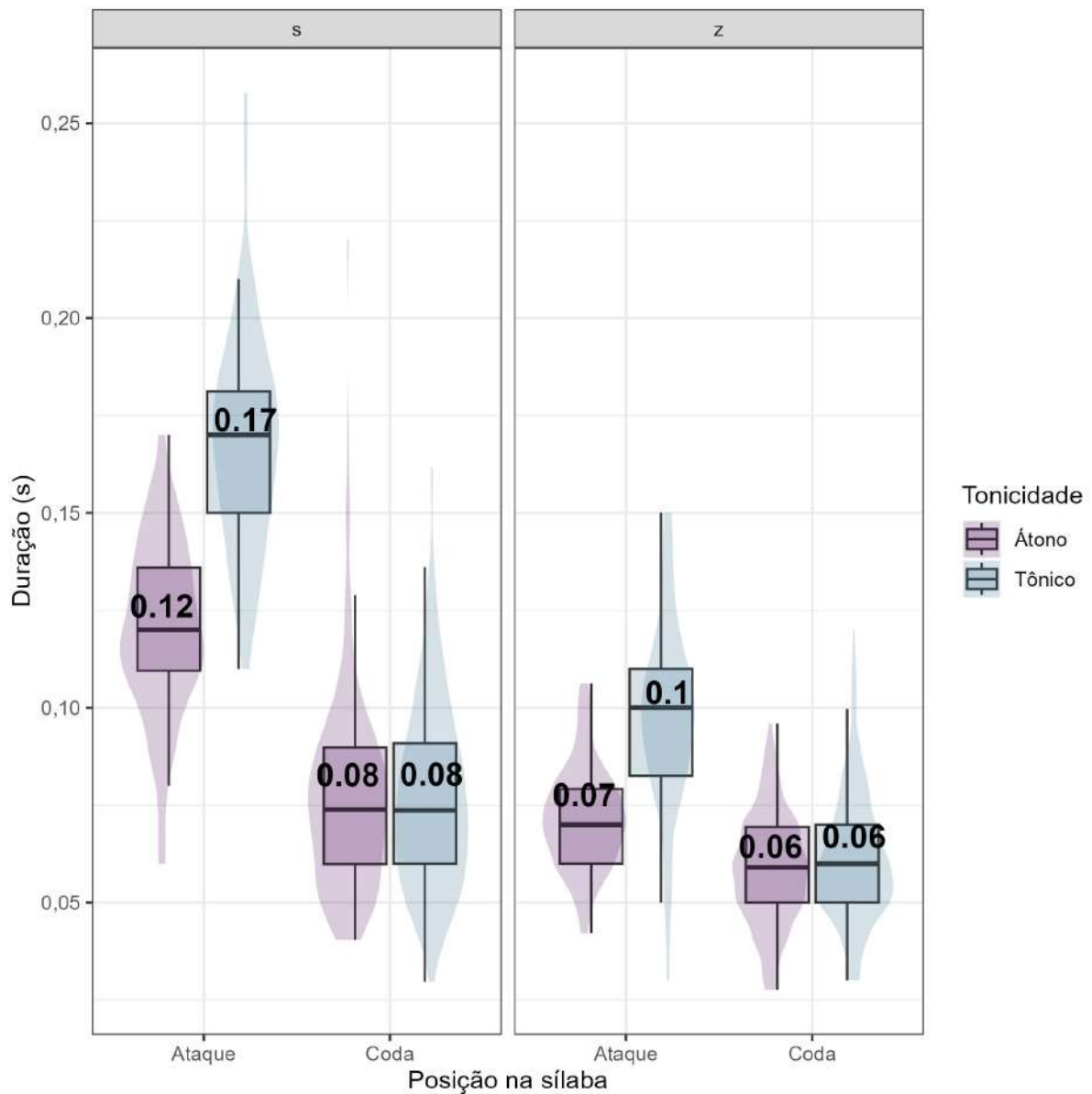


Figura 43 – Duração das duas realizações de /s/ de acordo com a tonicidade e a posição na sílaba.

Tabela 29 – Resultados do modelo misto para razão de duração da coda silábica de acordo com a sua posição na palavra.

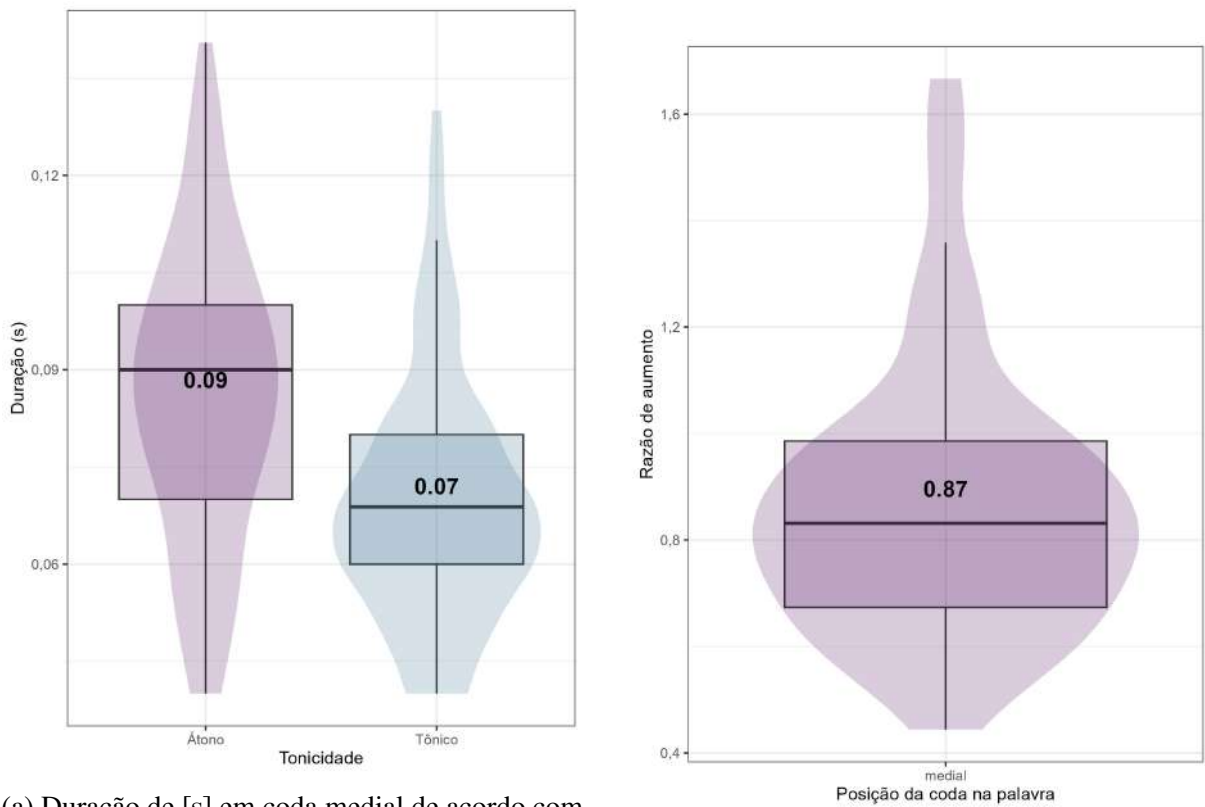
Efeito Fixo	Estimativa	Erro Padrão	df	Valor-t
Intercept (Medial)	0,87460	0,03962	36,48	22,07 ***
Coda final	0,23393	0,02899	708,79	8,07 ***

0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 ',' 0,1 ' ' 1

Tabela 30 – Resultados dos Contrastes de Tukey (Método de Holm).

Comparação	Estimativa	Erro Padrão	Valor-z	Valor-p ajust.
final - medial	0,23393	0,02899	8,07	$6,66 \times 10^{-16}$ ***

0 '\*\*\*' 0,001 '\*\*' 0,01 '\*' 0,05 ',' 0,1 ' ' 1



(a) Duração de [s] em coda medial de acordo com a tonicidade.

(b) Razão de alongamento do [s] em coda medial.

Figura 44 – Comparação entre a duração e a razão de duração do [s] em coda medial.

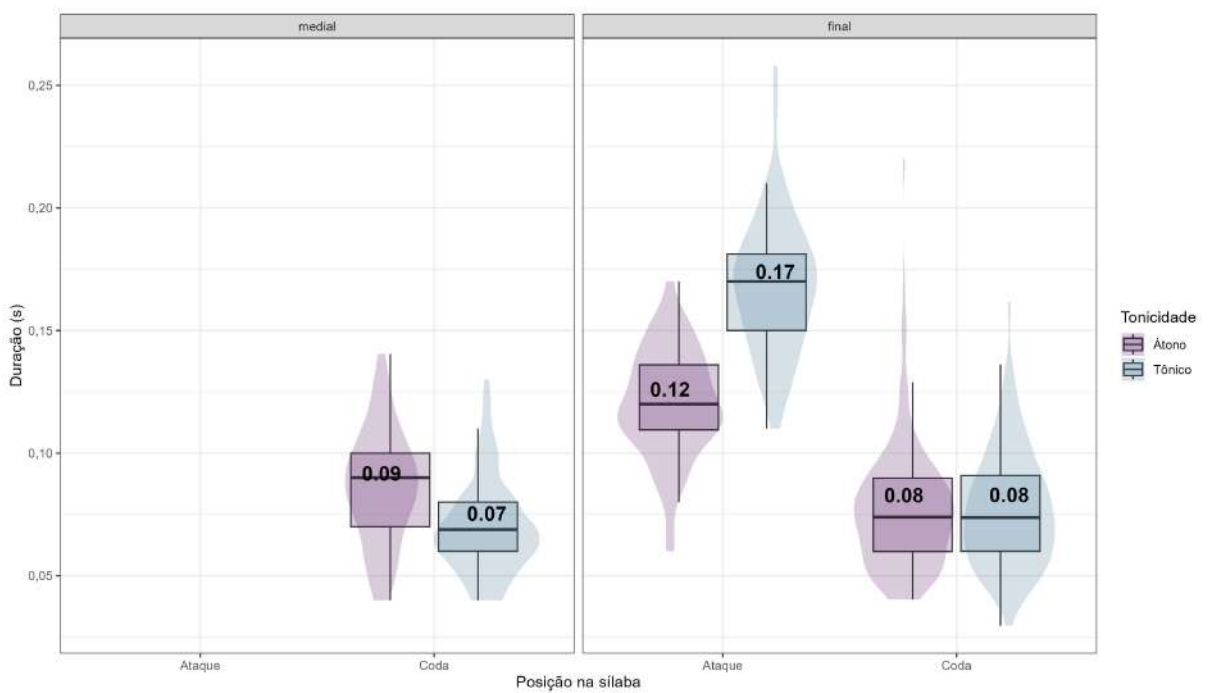


Figura 45 – Duração de [s] de acordo com a posição da coda na palavra, posição do segmento na sílaba e tonicidade.

## 5 Discussão

Os resultados obtidos corroboram a hipótese inicial deste estudo, que sugeria que a complexidade silábica e a posição segmental afetam a manifestação do acento de palavra no PB. As diferenças de duração observadas entre sílabas tônicas e átonas variam conforme a complexidade da sílaba e a posição dos segmentos.

Os dados indicam que a teoria da Fonologia Articulatória e o conceito de não-alinhamento Centro-C (BYRD, 1995) em codas silábicas são compatíveis com os resultados obtidos, sugerindo que a coda silábica, de fato, não contribui para o reajuste gestual. O comportamento da coda, com a sua aparente ressilabificação para o ataque da próxima sílaba e o seu consequente alongamento, foi um resultado já encontrado por Stetson (1951), e que permanece coerente dentro do (re)alinhamento de gestos dentro da sílaba. Por outro lado, o comportamento similar entre consoantes que iniciam ataques simples e complexos parece indicar que, nesses casos, não há reajuste gestual perceptível.

A seguir, serão discutidos os efeitos do acento de palavra na sílaba como um todo e nos seus vários constituintes.

### 5.1 A sílaba

As medidas de duração para toda a unidade silábica destacaram alguns efeitos do tipo de molde da sílaba e de seus segmentos internos: resumidamente, quanto mais complexo o molde, menor é o alongamento sofrido pela sílaba tônica em relação ao seu par átono.

Dessa forma, o molde CV é o que mais passa pelo processo de alongamento devido ao acento de palavra, seguido dos moldes CVC e CCV. Sílabas CCVC são as que menos possuem suas durações afetadas. Uma explicação para esse cenário somente pode ser realmente construída se for levado em conta aquilo que ocorre com a vogal /a/ e as demais consoantes dentro da sílaba.

Como observado nos resultados nas Subseções 4.2.1 e 4.2.2, todos os segmentos tônicos – com possível exceção do tepe<sup>1</sup> – apresentam maior duração do que seus pares átonos. No entanto, o núcleo da sílaba se destaca dos demais constituintes, medindo aproximadamente o dobro sob efeito do acento de palavra, enquanto as consoantes exibem valores de alongamento menores<sup>2</sup>.

Então, os resultados obtidos para a unidade silábica podem ser compreendidos como

<sup>1</sup> Ver discussão na Subseção 5.2.2.2.

<sup>2</sup> É importante destacar que este trabalho não assume que o efeito do acento de palavra sobre as consoantes, por causa do fato acima, seja desprezível. Pelo contrário, argumenta-se que a consideração da duração consonantal é fundamental, uma vez que análises anteriores frequentemente desconsideraram esse aspecto.

uma redistribuição do alongamento proporcionado pelo acento de palavra. Assim, o impacto do acento na duração total da sílaba é suavizado pelo peso proporcionalmente menor das consoantes, resultando em uma alteração menos aparente da duração. À medida que a complexidade do molde silábico aumenta (de CV para CCV, CVC e CCVC), a vogal passa a constituir uma proporção menor da estrutura total da sílaba, o que reduz seu impacto relativo na duração da unidade como um todo.

Uma hipótese para os resultados esperados da duração da sílaba não foi apresentada, pois o foco deste trabalho é analisar reestruturações temporais internas às sílabas. Essas reestruturações são invisíveis ou, no mínimo, inexatas quando se mede a duração da sílaba inteira sem subespecificações. No entanto, os resultados obtidos estão alinhados com o esperado. Embora poucos estudos analisem individualmente a duração dos constituintes silábicos sob efeito de acento, há um consenso de que as vogais, quando comparadas às consoantes, são mais afetadas por mudanças temporais. Isso ocorre porque as vogais podem ser sustentadas continuamente por períodos mais longos ou ter sua duração controlada pelo falante, dentro dos limites do fluxo de ar disponível para a fonação. Esse entendimento sugere que os efeitos do acento de palavra sobre a duração total da sílaba dependem da proporção que a vogal (ou núcleo) ocupa na estrutura silábica. Quanto maior for essa participação relativa, maior será o impacto do acento de palavra na duração da sílaba como um todo.

## 5.2 Os segmentos

### 5.2.1 O núcleo da sílaba: /a/

Os resultados obtidos a partir da medição de duração do núcleo das sílabas apontaram a um evidente alongamento em contexto tônico dessa parte da sílaba, assim como se hipotetizou. Pelo fato de a vogal ter sido o maior alvo de medições dos estudos referentes ao acento de palavra em diversas línguas do mundo, o resultado encontrado aqui era bastante esperado. Como se esperava, não houve efeito do molde silábico em seu alongamento, pelo papel de “âncora” rítmica.

Esses achados indicam que o núcleo da sílaba apresenta uma reestruturação temporal estável, independentemente das alterações no entorno. Enquanto as consoantes ao redor parecem ser “ancoradas” ao núcleo e passam por ajustes temporais para manter uma certa coordenação articulatória específica (efeito Centro-C), o núcleo mantém sua duração inalterada. Esse comportamento sugere que, embora o núcleo atue como um ponto de referência articulatório para os gestos consonantais, ele próprio é menos influenciado pelas reconfigurações temporais que ocorrem na estrutura silábica ao seu redor.

## 5.2.2 As partes não nucleares da sílaba

### 5.2.2.1 As consoantes em ataque simples ou ramificado

A hipótese de que haveria o alongamento das consoantes em primeira posição de ataque se confirmou e a codificação de mais variáveis durante as avaliações estatísticas apontou que o efeito do acento de palavra é sensível, não apenas ao molde, mas também ao vozeamento da consoante e ao seu modo de articulação. De maneira geral, as consoantes em posição C1 que mais se alongaram foram aquelas com molde CV, com modo de articulação fricativa e com o vozeamento presente.

O fato de a consoante iniciar um ataque sozinha ou ao lado do tepe, i.e. em ataque simples ou complexo, não foi relevante para afetar o seu alongamento. Isso parece apontar ao fato de que, embora em ataques complexos ocorra o realinhamento das consoantes em torno do seu ponto médio (efeito de Centro-C, apresentado na Subseção 2.3.3), a maneira com a qual cada segmento se alonga individualmente não é afetada.

O alongamento maior das consoantes fricativas pode estar relacionado à sua natureza acústico-articulatória contínua, que propicia a produção do som de maneira quase ininterrupta. A relação com o vozeamento, que faz as consoantes se alongarem ligeiramente mais, por outro lado, embora tenha sido apontada como estatisticamente significativa, não parece ser relevante linguisticamente.

### 5.2.2.2 A consoante em ramificação de ataque: [r]

A hipótese sobre o não alongamento de [r], a princípio, não se confirmou. Conforme apresentado, a presença de um efeito de alongamento no tepe devido ao acento de palavra era de certa forma inesperado devido à natureza desse som: rápido e preciso (LADEFOGED; MADDIESON, 1996). No entanto, assim como foi descrito na Subseção 3.2.3.2, por vezes estaria presente junto ao tepe um segmento vocálico, ao que se chamou de vogal de apoio (VA). A princípio, atribui-se o alongamento observado a esse som vocálico emergente, devido à evidente facilidade que vogais têm de se alongar.

Uma melhor investigação sobre este caso, porém, ainda há de ser feita, uma vez que a medição separada da VA e do momento de oclusão do tepe não foi realizada no presente trabalho. Entende-se que a avaliação do alongamento desse trecho vocálico de um som consonantal é pertinente para os estudos das propriedades temporais de sons complexos e/ou não contínuos e da manifestação acústica do acento de palavra.

### 5.2.2.3 A consoante em coda: /s/

Os resultados das medições de /s/ em coda silábica, num primeiro momento, revelam que o efeito do acento de palavra sobre a duração se manifesta de maneira distinta dependendo da posição do segmento na sílaba. Enquanto segmentos em posição de ataque apresentam um

alongamento significativo em sílabas acentuadas, os segmentos em coda aparentemente não demonstram esse comportamento da maneira como se esperava. Pode-se pensar que o padrão observado corrobora a hipótese inicial de que as codas, devido à sua coordenação anti-fase em relação ao núcleo silábico, seriam menos suscetíveis ao alongamento induzido pelo acento, conforme argumentado anteriormente com base no efeito Centro-C descrito pela Fonologia Articulatória.

Uma nova interpretação se faz também possível a partir da observação dos dados de [s] em coda em meio de palavra. Uma análise mais detalhada, comparando codas em posições finais e mediais de palavras, revela que as codas em sílabas finais, mesmo quando acentuadas, são mais curtas do que suas contrapartes não acentuadas em posições mediais de palavra. Esse resultado, embora inesperado, pode estar relacionado à ressilabificação homossilábica em grupos consonantais na posição medial de palavras, como sugerido pela Figura 46.

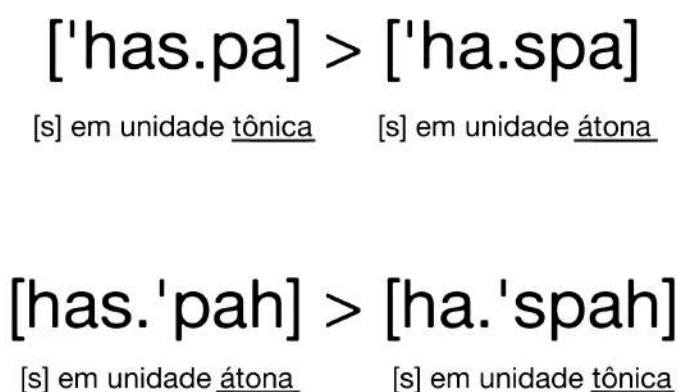


Figura 46 – Processo de ressilabificação exemplificado por ‘raspa’ e ‘raspar’.

Essa análise está de acordo com os resultados de [Stetson \(1951\)](#), em que o autor

procurou construir uma noção de sílaba compatível com a sua visão da fala como movimento audível. Mostrou experimentalmente que, **na pronúncia repetida e acelerada** de palavras iniciadas por vogal e terminadas por consoante, a coda final subitamente muda de posição e passa a ocupar o ataque inicial. **Assim, em português, ‘ás, ás, ás, ás’ transforma-se em ‘Sá, Sá, Sá, Sá’ [...].** ([ALBANO, 2020](#), p. 73, grifos nossos)

Vendo por esse lado, então, o [s] sofre, sim, o efeito do acento de palavra, mas somente daquele que recai sobre a sílaba seguinte. Esse entendimento não refuta a interpretação baseada no efeito Centro-C, já que, por agora pertencer ao ataque de uma sílaba, os gestos associados a [s] podem ser rearranjados temporalmente. Essa reorganização dos gestos articulatórios pode ser compreendida dentro de uma perspectiva dinâmica, em que mudanças na taxa de fala influenciam a estabilidade dos gestos de coda. [Albano \(2020\)](#) explica que

*[se trata] do fenômeno, largamente observado no mundo físico, da mudança de regime dinâmico a partir de um nível crítico de instabilidade. Assim, a aceleração crescente da taxa de elocução leva os gestos de coda a uma instabilidade motora crítica, causando a sua atração automática para o regime dinâmico mais estável e levando-os a entrar em fase com a vogal seguinte. (ALBANO, 2020, p. 76, grifos nossos)*

Quanto a isso, cabe avaliar o efeito da taxa de fala de cada participante com essa aparente reestruturação da sílaba com coda, tarefa que fica para futuros trabalhos. Futuras análises devem analisar o efeito da taxa de articulação na mudança de coordenação da coda.

Outro ponto que parece reforçar isso, e que deve ser lembrado, é o fato de que a palavra seguinte às sílabas finais medidas no experimento é ‘baixinho’, que se inicia por uma sílaba não acentuada. Esse cenário está de acordo com a linha de raciocínio construída aqui, por confirmar o motivo pelo qual a coda final não é alongada pelo acento de palavra: a sílaba seguinte, da qual o som reesilabificado faz parte, é átona e não promove o alongamento esperado. Por outro lado, é muito importante o apontamento de que, em fase bastante inicial da presente pesquisa, em que a palavra seguinte da frase-veículo era iniciada por uma sílaba com acento primário, a medição da duração do /s/ em coda levou aos mesmos resultados, ainda que com o número de participantes bem inferior (cinco). Por esse fato, é justo imaginar a possibilidade de a ressilabificação proposta estar restrita à própria palavra, i.e. o processo pode ser interrompido por fronteiras prosódicas mais altas que a sílaba ( $\sigma$ ) e que o pé ( $\Sigma$ ), barrado por fronteira de palavra (#).

## Os encontros-sC

Um ponto digno de nota é o fato de o processo de ressilabificação, como pode ser visto na Figura 46, resultar em uma sílaba que vai contra o Princípio do Sequenciamento de Soância<sup>3</sup>. Em outras palavras, não seria esperado que uma sílaba tivesse uma sequência sonora como ilustrado na Figura 47, onde se observa uma descida de soância de /s/ para /p/ antes de chegar ao seu pico no núcleo.

Essa situação provoca indagações quanto a se de fato está havendo um processo de ressilabificação nos dados analisados, ou, por outro lado, se o princípio da Teoria de Soância é mesmo relevante para as sílabas do PB. Uma alternativa é considerar que, enquanto o Princípio da Sequência de Soância se aplicaria a uma sílaba, a manifestação do acento de palavra teria como unidade de aplicação alguma outra que corresponderia ao *offset*, i.e. o fim, de uma vogal até o *offset* da vogal seguinte.

A sequência sC<sup>4</sup> (uma fricativa sibilante seguida de outra consoante, e também referida como *sC-cluster*) é problemática para a Teoria de Soância porque ela frequentemente viola o princípio central dessa teoria: a escala de soância. A teoria postula que, dentro de uma sílaba,

<sup>3</sup> Conceito já apresentado e discutido na Subseção 2.1.2.

<sup>4</sup> Ver Goad (2011) para uma discussão sobre a representação desse tipo de encontro.

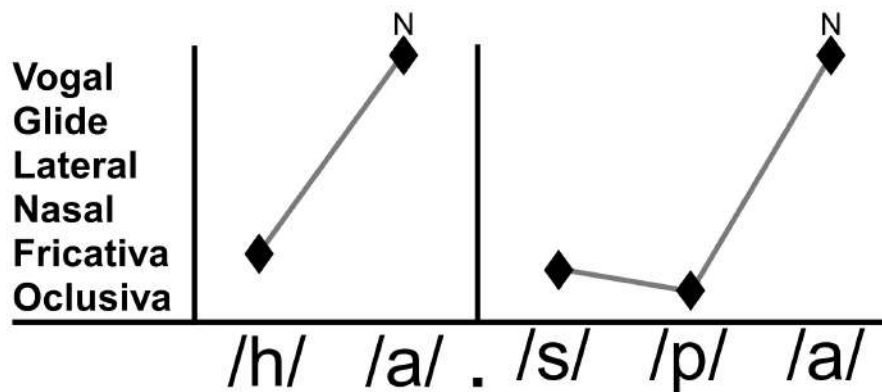


Figura 47 – Representação da escala de soância nas sílabas de <raspa> após o processo de ressilabificação proposto. N = núcleo da sílaba.

os segmentos devem ser organizados de modo que a soância aumente em direção ao núcleo da sílaba (geralmente uma vogal). No entanto, em muitos casos, a sequência sC não segue essa progressão natural.

Por exemplo, em encontros como [sp], [st] ou [sk], o segmento [s] possui menor soância do que a consoante seguinte (como [p], [t] ou [k]), mas aparece na posição inicial do encontro. Isso contraria a expectativa da teoria, segundo a qual o som mais sonoro (a consoante seguinte) deveria aparecer primeiro. Essa violação sugere que [s] em sC pode ser tratada de forma diferente na estrutura silábica, como é apresentado na Figura 48.

Encontros-sC já foram alvo de avaliações de soância e apresentam desafios para as teorias de representação fonológica. Esses *clusters* parecem se comportar como coda-ataque, mas também podem ocorrer no início das palavras, onde seu comportamento difere dos ataques ramificados. A análise de [s] nesses encontros consonantais iniciais varia, sendo a análise de coda a mais satisfatória segundo Polgárdi (2022), embora enfrente o problema de garantir que o núcleo vazio anterior permaneça silencioso, cf. a Figura 48.

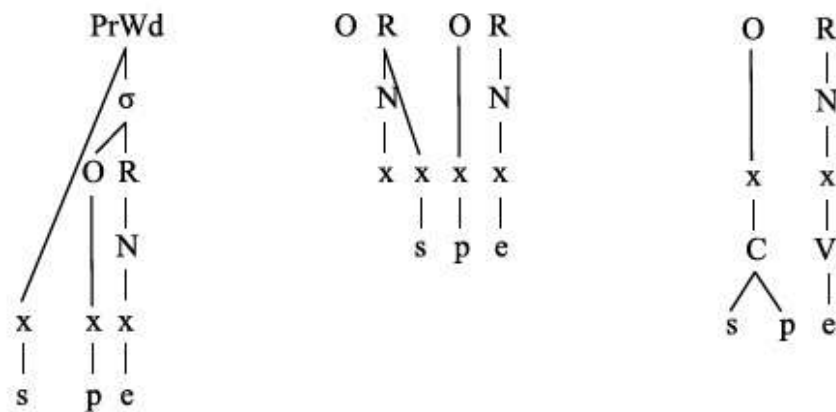


Figura 48 – Análises encontradas ao longo da literatura para explicar encontros-sC iniciais. Da esquerda para a direita: [s] em apêndice à sílaba; [s] em coda de sílaba de núcleo não-realizado; e [sp] como um segmento complexo. Fonte: retirado de Polgárdi (2022, p. 2).

Tilsen (2019) cita como encontros-sC também são críticos na literatura para tratar do efeito Centro-C proposto pela Fonologia Articulatória, apresentando um comportamento diferente de outros encontros consonantais:

*Hermes et al. (2013) found evidence for simplex organization of /sC/ sequences in Italian, but complex organization of other clusters, which suggests that control of timing can be segment-specific. Hence there appear to be two control strategies available for production of word-initial consonant clusters, one in which the consonants are organized globally into an onset, the other in which only the pre-vocalic consonant is organized with the vowel.*

Os resultados do presente trabalho fornecem mais meios para que se amplie a discussão em torno de encontros-sC, especificamente nos casos de encontros heterossilábicos, evidenciando como nesses casos o [s] é afetado por efeitos de alongamento do acento da sílaba seguinte, i.e. de C. Tendo isso em consideração, é curioso o fato de encontros-sC tautossilábicos em começo de sílaba serem separados em duas sílabas diferentes, uma das quais terá um núcleo nulo, enquanto, em casos de encontros heterossilábicos, os segmentos seriam juntados. Mais testes e trabalhos sobre este tema são necessários para traçar uma relação entre os cenários dispostos anteriormente, porém já é possível que se aponte direções por onde se começar a questionar análises tradicionais de encontros-sC: será importante entender e considerar o papel da tonicidade na estruturação silábica durante a análise.

### O Inter Perceptual Center Group (IPCG)

De maneira semelhante às implicações desse tipo de encontro consonantal para a estrutura silábica de uma língua, observando um fenômeno similar, Barbosa (1996) propõe a existência de uma outra unidade rítmica (relevante especificamente para a percepção) dessas sequências sonoras, o *Inter Perceptual Center Group (IPCG)*.

De acordo com o autor, a duração de consoantes em encontros como [st] não deve ser analisada isoladamente, mas como parte de uma unidade perceptual agrupada. A proposta surge ao observar que, nos dados de um falante paulista, em palavras como ‘passa’ ([ˈpas.sa]) e ‘pasta’ ([ˈpas.ta]), a duração do [s] em posição de ataque (como em ‘passa’) é maior do que a duração do [s] em coda de um encontro consonantal (como em ‘pasta’). No entanto, nessa visão, a soma das durações de [s] e [t] em ‘pasta’ equivaleria aproximadamente à duração do [s] em ‘passa’. Esse padrão sugere que, no caso de ‘pasta’, o [s] complementa a duração do [t], formando um agrupamento perceptual coerente. Seria interessante poder analisar o efeito do acento de palavra em uma comparação experimental semelhante (e.g. ‘passa’ vs. ‘pasta’ e ‘passar’ vs. ‘pastar’), e com um maior número de dados.

Os resultados do presente estudo apresentam paralelos com o padrão descrito, ao comparar a duração de [s] em combinações como [s] átono + consoante tônica (‘raspar’, ‘mascar’) e [s]ônico + consoante átona (‘raspa’, ‘masca’). Foi constatado que a duração de [s] em coda é sensível à tonicidade da sílaba seguinte, sugerindo uma interação prosódica entre o [s] e a sílaba tônica subsequente. Isso pode ser interpretado como uma extensão do conceito de IPCG, onde a interação duracional depende não apenas do agrupamento consonantal, mas também do contexto prosódico mais amplo. A continuidade desse tipo de estudo, com maior controle experimental, poderia aprofundar a compreensão de como o IPCG interage com fatores prosódicos como tonicidade e estrutura silábica.

Entende-se que, antes de traçar conclusões mais esclarecedoras e concretas, é necessária a elaboração de um estudo focado na análise do fenômeno aqui descrito, com um desenho experimental específico para explorar o caso.

## 6 Conclusão

O presente estudo se empenhou em avaliar a manifestação acústica, pela duração, do acento de palavra no português brasileiro. Foi realizada uma revisão da literatura relevante sobre a sílaba, o acento de palavra e a Fonologia Articulatória, destacando autores e trabalhos essenciais para a compreensão desses tópicos.

Foi realizado um experimento de produção com a gravação de leitura de centenas frases e de diversos falantes, a fim de construir uma pesquisa com boas práticas de estatística. O protocolo seguido foi criterioso, partindo do que já se sabe sobre o acento do nível da palavra, para que o estudo captasse bem os seus efeitos na fala dos participantes. A segmentação automatizada dos 180 minutos de gravações tornou possível trabalhar com um número tão grande e não usual de dados para estudos em fonologia. Ainda que com alguns fenômenos fonológicos aparecendo ao longo do *corpus* gravado, a análise ainda se fez bastante possível com o auxílio de ferramentas digitais, como *softwares* e *scripts*. A modelagem estatística das mais de 6.000 observações e 22 variáveis também foi alvo de inúmeras revisões ao longo de toda a pesquisa, e foi alvo de alto grau de cuidado, por ser também um ponto de empenho desta dissertação.

Um ponto caro para esta pesquisa foi a avaliação de segmentos e, conseqüentemente, de estruturas silábicas que foram e ainda são desconsiderados nos estudos fonológicos, principalmente naqueles que compartilham do complexo objetivo de medir a materialização do acento de palavra. Procurou-se aqui mostrar que nem sempre aquilo que acontece com a vogal é o que se reflete para a toda sílaba; e nem tudo aquilo que acontece com a sílaba CV vale também para os outros tipos silábicos. O estudo baseado em áudios de falantes do português brasileiro permitiu que essa proposta fosse alcançada, aproveitando a rica estrutura silábica da língua. Além disso, foram utilizadas ferramentas digitais disponíveis atualmente (e.g. *EasyAlign* e *RStudio*), com o objetivo de garantir que barreiras tecnológicas não impedissem a investigação das hipóteses de forma competente.

O presente trabalho buscou, ainda, se embasar teoricamente em um modelo relativamente inovador que propiciava ferramentas para elaborar pensamentos em torno da sílaba, unidade fonológica tão essencial para a realização desta pesquisa. Também, a maneira como a Fonologia Articulatória trata do *tempo* em sua abordagem faz com que um estudo focado na medição da duração, unidade temporal, ganhe muito quanto às ferramentas teóricas propostas. Um melhor entendimento de como o acento de palavra afeta unidades linguísticas abre caminho, ademais, para o avanço da modelagem prosódica da FA, que é tão recente. Uma teoria em fonologia deve estar preparada para lidar com fenômenos em um nível global, e não *language-specific*.

Vários pontos, que seriam dignos de um estudo específico, surgiram ao longo desta pesquisa, e não puderam ser analisados como deveriam. Nesse sentido, o aparente alongamento

da vogal de apoio do [r] e a interação da duração dos elementos em coda com a sílaba seguinte, e não com a sua própria, foram resultados que só poderiam ter sido encontrados sob uma análise que fosse além de sílabas CV e de núcleo silábicos. O caso do tepe abre questões importantes sobre a estrutura gestual desse som, assim como a natureza da sua vogal de apoio, e perguntas podem surgir, tais como: qual é a relação da Vogal de Apoio com a vogal do núcleo da sílaba? A Vogal de Apoio realmente deveria ser atribuída a uma parte do tepe? Já o comportamento da coda silábica faz aparecer novas formas de identificar processos de ressilabificação, fenômeno tão caro a qualquer teoria fonológica.

Uma limitação desta pesquisa foi o uso da fala de laboratório para a avaliação do fenômeno estudado. Esse tipo de dado impõe restrições quanto à naturalidade da produção. A leitura de sentenças pode induzir um padrão entoacional menos espontâneo e potencialmente enviesar a manifestação acústica do acento de palavra em comparação com a fala espontânea (BECKMAN, 1997; WARNER et al., 2012). Porém, seu uso se justifica por oferecer um controle rigoroso das variáveis experimentais, possibilitar uma análise detalhada dos efeitos acústicos do acento de palavra e viabilizar a análise de um número de dados relativamente extenso. Xu (2010) argumenta a favor do uso da fala controlada, discutindo seus pontos fortes e como contornar seus problemas. Estudos futuros poderão complementar essa abordagem com dados de fala semi-espontânea ou conversacional, a fim de verificar até que ponto os padrões identificados aqui se fazem presentes em contextos mais naturais de fala.

Um processo de segmentação ao nível dos gestos também beneficiaria o presente trabalho. Devido ao número de dados analisados e ao tempo disponível para este estudo, somente os fones foram segmentados e anotados, o que impôs certas limitações na interpretação dos resultados. A segmentação baseada em gestos permitiria uma análise mais detalhada da temporalidade dos articuladores e de suas sobreposições, possibilitando uma compreensão mais precisa da relação entre os gestos subjacentes e os padrões duracionais do acento de palavra.

Uma adição valiosa ao estudo seria também o uso de dados articulatórios, que têm se tornado mais acessíveis apenas nos últimos anos, mas que ainda não estão disponíveis na Faculdade de Letras da UFMG. Embora tenham suas ressalvas, técnicas como ultrassonografia de língua, eletromagnetometria e ressonância magnética dinâmica poderiam fornecer informações mais detalhadas sobre os gestos articulatórios subjacentes aos padrões acústicos observados, permitindo uma análise mais aprofundada da interação entre gestos e prosódia. A incorporação desses métodos possibilitaria testar de maneira mais direta hipóteses relacionadas à estrutura gestual da sílaba e ao impacto do acento de palavra nos diferentes níveis da fala, assim como diversos estudos baseados na Fonologia Articulatória têm feito.

Com um número maior de participantes e maior variabilidade de dados, futuras pesquisas poderão explorar mais a fundo essas questões e confirmar se os padrões observados se mantêm em outras estruturas silábicas e com diferentes tipos de consoantes e vogais. Mais avaliações com outras línguas também poderão fazer surgir um diálogo científico entre os trabalhos, de

maneira a fazer caminhar a literatura fonológica sobre o acento de palavra e prosódia no geral.

# Referências

- ABAURRE, M. B.; PAGOTTO, E. Nasalização vocálica no português do Brasil. In: KOCH, I. V. (Ed.). *Gramática do português falado*. Campinas: Editora da Unicamp, 1996. VI, p. 495–526.
- ABERCROMBIE, D. Stress and some other terms. *Work in Progress*, v. 9, p. 51–53, 1976. [Reprinted in D. Abercrombie, 1991. *Fifty years in phonetics*. Edinburgh: Edinburgh University Press.].
- ABRAMSON, A.; LISKER, L. A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, v. 20, p. 384–422, 07 1964.
- ABRAMSON, A. S.; REO, N. Distinctive vowel length: duration vs. spectrum in Thai. *Journal of Phonetics*, v. 18, n. 2, p. 79–92, 1990. ISSN 0095-4470. Linguistic Approaches to Phonetics Papers presented in Honor of J.C. Catford. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S009544701930395X>>.
- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v. 19, n. 6, p. 716–723, 1974.
- AKHVLEDIANI, G. S. *Zogadi ponet'ik'is sapudzvlebi*. Tbilisi: Tbilisi State University Publishing, 1949.
- ALBANO, E. O português brasileiro e as controvérsias da fonética atual: pelo aperfeiçoamento da fonologia articulatória. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, Pontifícia Universidade Católica de São Paulo - PUC-SP, v. 15, n. spe, p. 23–50, 1999. ISSN 0102-4450. Disponível em: <<https://doi.org/10.1590/S0102-44501999000300002>>.
- ALBANO, E. *O gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro*. Campinas: Mercado de Letras, 2001.
- ALBANO, E. *O gesto audível: fonologia como pragmática*. Cortez Editora, 2020. ISBN 9786555550030. Disponível em: <<https://books.google.com.br/books?id=KdjkdWAAQBAJ>>.
- ANDERSON, J. Syllabic or non-syllabic phonology? *Journal of Linguistics*, v. 5, n. 1, p. 136–142, 1969.
- AUDACITY Team. *Audacity 3.4.2*. 2023. Acesso: 19/02/2024. Disponível em: <<https://www.audacityteam.org/>>.
- BALENO, G.; JANZEN, J.; YODER, B. Acoustic correlates of word stress in Haisla. In: REISINGER, D. K. E. et al. (Ed.). *Papers for the International Conference on Salish and Neighbouring Languages 59*. Vancouver, BC: UBC WPL, 2024.
- BARBOSA, P. At least two macrorhythmic units are necessary for modeling Brazilian Portuguese duration. In: *Proceedings of the First ETRW on Speech Production Modeling*. [S.l.: s.n.], 1996. p. 85–88.
- BARBOSA, P.; ERIKSSON, A.; ÅKESSON, J. Cross-linguistic similarities and differences of lexical stress realisation in Swedish and Brazilian Portuguese. In: ASU, E. L.; LIPPUS, P. (Ed.). *Nordic Prosody. Proceedings of the XIth conference, Tartu 2012*. Frankfurt am Main: Peter Lang, 2012. p. 97–106.

- BARBOSA, P.; ERIKSSON, A.; ÅKESSON, J. On the robustness of some acoustic parameters for signalling word stress across styles in brazilian portuguese. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, p. 282–286, 01 2013.
- BARBOSA, P. A.; MADUREIRA, S. *Manual de fonética acústica experimental: aplicações a dados do português*. São Paulo: Cortez, 2015. 591 p. ISBN 978-85-249-2421-7.
- BARNES, J. *Positional neutralization: A phonologization approach to typological patterns*. Tese (Doutorado) — University of California, Berkeley, 2002.
- BARNES, J. *Phonetics and Phonology in Russian Unstressed Vowel Reduction: A Study in Hyperarticulation*. Boston, MA: [s.n.], 2007.
- BECKMAN, M.; EDWARDS, J.; FLETCHER, J. Prosodic structure and tempo in a sonority model of articulatory dynamics. In: \_\_\_\_\_. *Gesture, Segment, Prosody*. [S.l.]: Cambridge University Press, 1992. (Papers in Laboratory Phonology), p. 68–89.
- BECKMAN, M. E. *Stress and Non-Stress Accent*. Dordrecht: Foris, 1986. ISBN 90-6765-243-1.
- BECKMAN, M. E. A typology of spontaneous speech. In: SAGISAKA, Y.; CAMPBELL, N.; HIGUCHI, N. (Ed.). *Computing Prosody: Computational Models for Processing Spontaneous Speech*. New York: Springer Verlag, 1997. p. 7–26.
- BECKMAN, M. E.; EDWARDS, J. Articulatory evidence for differentiating stress categories. In: *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*. [S.l.]: Cambridge University Press, 1994. p. 7–33.
- BISOL, L. Sândi externo: O processo e a variação. In: KATO, M. (Ed.). *Gramática do Português Falado. Volume V: Convergências*. 2ª ed.. ed. Campinas, SP: Editora da UNICAMP, 2002.
- BISOL, L. *Introdução a estudos de fonologia do português brasileiro*. 5. ed. Porto Alegre: ediPUCRS, 2010.
- BISOL, L. A simetria no sistema vocálico do português brasileiro. *Linguística*, Porto, Portugal, v. 5, 2010.
- BLEVINS, J. The syllable in phonological theory. In: GOLDSMITH, J. (Ed.). *The handbook of phonological theory*. Cambridge, MA: Blackwell, 1995. p. 206–244.
- BOERSMA, P.; WEENINK, D. *Praat: doing phonetics by computer*. [S.l.], 2023. [Programa de computador] Versão 6.3.10, acesso em 3 maio 2023 em <<http://www.praat.org/>>.
- BOLINGER, D. Pitch accent and sentence rhythm. *Forms of English: Accent, morpheme, order*, v. 1, 1965.
- BORISE, L. Disentangling word stress and phrasal prosody: A view from georgian. *Phonological Data and Analysis*, v. 5, n. 1, p. 1–37, 2023.
- BROSCHES, C. de. *Traité de la formation mécanique des langues, et de principes physiques de l'étymologie*. Paris: Chez Saillant, Vincent, Desaint, 1765.
- BROWMAN, C. P.; GOLDSTEIN, L. M. Towards an articulatory phonology. *Phonology*, v. 3, n. 1, p. 219–252, 1986.

- BROWMAN, C. P.; GOLDSTEIN, L. M. Some notes on syllable structure in articulatory phonology. *Phonetica*, v. 45, n. 2-4, p. 140–155, 1988.
- BROWMAN, C. P.; GOLDSTEIN, L. M. Articulatory gestures as phonological units. *Phonology*, v. 6, n. 2, p. 201–251, 1989.
- BROWMAN, C. P.; GOLDSTEIN, L. M. Tiers in articulatory phonology, with some implications for casual speech. In: KINGSTON, J.; BECKMAN, M. (Ed.). *Papers in Laboratory Phonology I: Between the grammar and physics of speech*. Cambridge: Cambridge University Press, 1990. p. 341–376.
- BURRONI, F.; TILSEN, S. The online effect of clash is durational lengthening, not prominence shift: Evidence from Italian. *Journal of Phonetics*, v. 91, p. 101124, 2022. ISSN 0095-4470. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0095447021000991>>.
- BYRD, D. C-centers revisited. *Phonetica*, v. 52, p. 285–306, 1995.
- BYRD, D.; SALTZMAN, E. Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, v. 26, n. 2, p. 173–199, 1998. ISSN 0095-4470. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0095447098900717>>.
- BYRD, D.; SALTZMAN, E. The elastic phrase: modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, v. 31, n. 2, p. 149–180, 2003. ISSN 0095-4470. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0095447002000852>>.
- CANTONI, M. M. *O Acento no Português Brasileiro: Uma Abordagem Experimental*. Tese (Tese de Doutorado) — Universidade Federal de Minas Gerais, 2013.
- CANTONI, M. M. A epêntese no português brasileiro em uma perspectiva multirrepresentacional. *Gragoatá*, v. 20, n. 38, jun. 2015. Disponível em: <<https://periodicos.uff.br/gragoata/article/view/33308>>.
- CANTONI, M. M.; PRADO, V. A. M. Word stress effects on syllable onset in Brazilian Portuguese. In: JAURENA, M.; METZLER, Z. (Ed.). *Proceedings of the 35th Western Conference on Linguistics*. Fresno, CA: California State University, Fresno, 2023. (WECOL, v. 29), p. 1. Hosted by Department of Linguistics, California State University, Fresno.
- CHO, T.; MCQUEEN, J. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, v. 33, p. 121–157, 2005.
- CLEMENTS, G. N. The role of the sonority cycle in core syllabification. In: KINGSTON, J.; BECKMAN, M. (Ed.). *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, 1990. p. 283–333.
- CLEMENTS, G. N.; KEYSER, S. J. *CV Phonology: A Generative Theory of the Syllable*. Cambridge, MA: MIT Press, 1983.
- COLE, J. et al. Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from radio news speech. *Journal of Phonetics*, v. 35, p. 180–209, 2007.
- COLLISCHONN, G. A epêntese vocálica no português do sul do Brasil. In: BISOL, L.; BRESCANCINI, C. (Ed.). *Fonologia e variação: recortes do português brasileiro (VARSUL)*. Porto Alegre: EDIPUCRS, 2002.

CROSSWHITE, K. Spectral tilt as a cue to word stress in polish, macedonian, and bulgarian. In: *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS-15)*. [S.l.: s.n.], 2003. p. 767–770.

CROSSWHITE, K. M. Vowel reduction. In: \_\_\_\_\_. *Phonetically Based Phonology*. [S.l.]: Cambridge University Press, 2004. p. 191–231.

CUTLER, A. Lexical stress. In: PISONI, D. B.; REMEZ, R. E. (Ed.). *The Handbook of Speech Perception*. Blackwell Pub, 2005. p. 264–289. Disponível em: <<https://doi.org/10.1002/9780470757024.ch11>>.

CUTLER, A.; JESSE, A. Word stress in speech perception. In: PARDO, J. S. et al. (Ed.). *The Handbook of Speech Perception*. 2. ed. John Wiley & Sons, 2021. Disponível em: <<https://doi.org/10.1002/9781119184096.ch9>>.

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

CÂMARA JÚNIOR, J. M. A acentuação e o vocábulo fonológico. In: *Estrutura da língua portuguesa*. 5. ed. Petrópolis: Vozes, 1975. p. 52–55.

CÂMARA JÚNIOR, J. M. *Estrutura da língua portuguesa*. 12. ed. Petrópolis: Editora Vozes, 1982.

DAVIS, S. *Topics in Syllable Phonology*. Tese (Doutorado) — University of Arizona, New York, 1985. Published in 1988.

DERWING, B.; EDDINGTON, D. The experimental investigation of syllable structure. *The Mental Lexicon*, v. 9, 11 2014.

DESCHAMPS, D. *Mecanismos nasais no português*. Dissertação (Dissertação de Mestrado) — Universidade Federal de Santa Catarina, Florianópolis, SC, 1976. 107 f.

D'IMPERIO, M.; ROSENTHALL, S. Phonetics and phonology of main stress in italian. *Phonology*, JSTOR, v. 16, n. 1, p. 1–28, 1999. Accessed 13 Feb. 2025. Disponível em: <<http://www.jstor.org/stable/4420141>>.

DOCHERTY, G. J. *The Timing of Voicing in British English Obstruents*. Berlin, New York: De Gruyter Mouton, 1992. ISBN 9783110872637. Disponível em: <<https://doi.org/10.1515/9783110872637>>.

DOGIL, G. The phonetic manifestation of word stress. In: HULST, H. van der (Ed.). *Word prosodic systems in the languages of Europe*. Berlin: Mouton de Gruyter, 1999. p. 273–334.

EDDINGTON, D. Spanish phonology and morphology: Experimental and quantitative perspectives. v. 53, p. 1–197, 12 2004.

EDWARDS, J.; BECKMAN, M. E.; FLETCHER, J. The articulatory kinematics of final lengthening. *The Journal of the Acoustical Society of America*, v. 89, n. 1, p. 369–382, 01 1991. ISSN 0001-4966. Disponível em: <<https://doi.org/10.1121/1.400674>>.

ERIKSSON, A. et al. The acoustics of lexical stress in italian as a function of stress level and speaking style. In: *Proceedings of Interspeech*. [S.l.: s.n.], 2016. v. 17, p. 1059–1063.

EVERETT, K. The acoustic correlates of stress in pirahã. *Journal of Amazonian Languages*, v. 1, n. 2, p. 104–162, 1998.

- FLETCHER, J. The prosody of speech: Timing and rhythm. In: HARDCASTLE, W. J.; LAVER, J.; GIBBON, F. E. (Ed.). *The Handbook of Phonetic Sciences*. United Kingdom: Wiley-Blackwell Publishing, 2010. p. 523–602.
- FOURAKIS, M. Tempo, stress, and vowel reduction in american english. *The Journal of the Acoustical Society of America*, v. 90, n. 4, p. 1816–1827, 1991.
- FRY, D. B. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, v. 27, p. 765–768, 1955.
- FRY, D. B. Experiments in the perception of stress. *Language and Speech*, v. 1, n. 2, p. 126–152, 1958. Disponível em: <<https://doi.org/10.1177/002383095800100207>>.
- FUDGE, E. Syllables. *Journal of Linguistics*, v. 5, p. 253–286, 1969.
- GOAD, H. The representation of sc clusters. In: \_\_\_\_\_. *The Blackwell Companion to Phonology*. John Wiley & Sons, Ltd, 2011. cap. 38, p. 1–26. ISBN 9781444335262. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/9781444335262.wbctp0038>>.
- GOEDEMANS, R.; HULST, H. van der. Weight-sensitive stress (v2020.3). In: DRYER, M. S.; HASPELMATH, M. (Ed.). *The World Atlas of Language Structures Online*. Zenodo, 2013. Disponível em: <<https://doi.org/10.5281/zenodo.7385533>>.
- GOEDEMANS, R.; ZANTEN, E. van. Stress and accent in indonesian. *LOT Occasional series*, v. 9, p. 35–62, 2007.
- GOFFMAN, L.; MALIN, C. Metrical effects on speech movements in children and adults. *Journal of Speech, Language, and Hearing Research*, v. 42, p. 1003–1015, 1999.
- GOLDMAN, J.-P. Easyalign: An automatic phonetic alignment tool under praat. In: *Interspeech'11, 12th Annual Conference of the International Speech Communication Association*. Firenze, Italy: [s.n.], 2011.
- GOLDSTEIN, L.; CHITORAN, I.; SELKIRK, E. Syllable structure as coupled oscillator modes: Evidence from georgian vs. tashlihyt berber. In: *Proceedings of the XVIth International Congress of Phonetic Sciences*. [S.l.: s.n.], 2007. p. 241–244.
- GORDON, M. Acoustic properties of primary and secondary word-level stress in estonian. In: *Poster presented at the 130th meeting of the Acoustical Society of America*. St. Louis: [s.n.], 1995.
- GORDON, M. Weight-sensitive stress. In: *Syllable weight: phonetics, phonology, typology*. New York: Routledge, 2006. cap. 4.
- GORDON, M. Stress: Phonotactic and phonetic evidence. In: OOSTENDORP COLIN J. EWEN, E. H. Marc van; RICE, K. (Ed.). *The Blackwell Companion to Phonology*. [S.l.]: John Wiley & Sons, 2011. p. 1–25.
- GORDON, M.; APPLEBAUM, A. Acoustic correlates of stress in turkish kabardian. *Journal of the International Phonetic Association*, v. 40, p. 35–58, 2010.
- GORDON, M.; HULST, H. van der. Word stress. In: GUSSENHOVEN, C.; CHEN, A. (Ed.). *The Handbook of Prosody*. Oxford: Oxford University Press, 2020. p. 66–77.

- GORDON, M.; ROETTGER, T. Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard*, v. 3, 08 2017.
- GUDAVA, T. Maxvilis adgilisatvis megrulši [stress placement in megrelian]. In: *Giorgi Akhvlediani Festschrift*. Tbilisi: Tbilisi University Press, 1969. p. 106–111.
- GUSSENHOVEN, C. *The Phonology of Tone and Intonation*. 1st. ed. Cambridge University Press, 2004. Disponível em: <<https://doi.org/10.1017/CBO9780511616983>>.
- HAIMAN, J. *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. Amsterdam: John Benjamins, 1980.
- HALL, N. Articulatory phonology. *Language and Linguistics Compass*, v. 4, n. 9, p. 818–830, 2010. Disponível em: <<https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/j.1749-818X.2010.00236.x>>.
- HALLE, M.; VERGNAUD, J.-R. *An Essay on Stress*. Cambridge, MA: The MIT Press, 1987.
- HARRIS, M. J. *Quantifying Speech Rhythms: Perception and Production Data in the Case of Spanish, Portuguese, and English*. Tese (PhD dissertation) — UC Santa Barbara, 2015.
- HAUGEN, E.; HALLE, M. et al. The syllable in linguistic description. In: HAUGEN, E.; HALLE, M. et al. (Ed.). *For Roman Jakobson*. [S.l.]: The Hague, 1956. p. 213–221.
- HAYES, B. Compensatory lengthening in moraic phonology. *Linguistic Inquiry*, v. 20, p. 253–306, 1989.
- HERMES, A.; MÜCKE, D.; GRICE, M. Gestural coordination of italian word-initial clusters: The case of 'impure s'. *Phonology*, v. 30, n. 1, p. 1 – 25, 2013. Cited by: 51. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84877014170&doi=10.1017%2fS095267571300002X&partnerID=40&md5=93c4d664b2fd8421f1333cbf35092e8b>>.
- HEUVEN, V. J. van. Acoustic correlates and perceptual cues of word and sentence stress: Towards a cross-linguistic perspective. In: GOEDEMAN, R. W. N.; HEINZ, J.; HULST, H. G. van der (Ed.). *The Study of Word Stress and Accent: Theories, Methods and Data*. Cambridge: Cambridge University Press, 2018. p. 15–59.
- HIRATA, Y. Effects of speaking rate on the vowel length distinction in japanese. *Journal of Phonetics*, v. 32, p. 565–589, 2004.
- HOCKETT, C. *A Manual of Phonology*. Bloomington: Indiana University Press, 1955.
- HONOROF, D. N.; BROWMAN, C. The center or edge: how are consonant clusters organised with respect to the vowel? In: ELENUS, K.; BRANDERUD, P. (Ed.). *Proceedings of the XIIIth International Congress of Phonetic Sciences, Vol.3*. Stockholm, Sweden: Congress Organisers at KTH and Stockholm University, 1995. p. 552–555.
- HULST, H. van der. Theories of the syllable. In: HULST, H. van der; RITTER, N. A. (Ed.). *The syllable: views & facts*. Berlin: Mouton de Gruyter, 1999. p. 13–52.
- HULST, H. van der. *Word Stress: Theoretical and Typological Issues*. Cambridge: Cambridge University Press, 2014.

- HURVICH, C. M.; TSAI, C.-L. Regression and time series model selection in small samples. *Biometrika*, v. 76, n. 2, p. 297–307, 06 1989. ISSN 0006-3444. Disponível em: <<https://doi.org/10.1093/biomet/76.2.297>>.
- HYMAN, L. Word-prosodic typology. *Phonology*, v. 23, p. 225 – 257, 08 2006.
- HYMAN, L. M. *A theory of phonological weight*. Dordrecht: Foris, 1985.
- ISKAROUS, K.; POUPLIER, M. Advancements of phonetics in the 21st century: A critical appraisal of time and space in articulatory phonology. *Journal of Phonetics*, Elsevier, 2022.
- JONG, K. de. Temporal constraints and characterising syllable structuring. In: LOCAL, J.; OGDEN, R.; TEMPLE, R. (Ed.). *Phonetic Interpretation: Papers in Laboratory Phonology VI*. Cambridge: Cambridge University Press, 2004. p. 253–268.
- KAHN, D. *Syllable-based Generalizations in English Phonology*. Indiana University Linguistics Club, 1976. (IU Linguistics Club). Disponível em: <<https://books.google.com.br/books?id=ELtZAAAAMAAJ>>.
- KANG, Y.; YOON, T.-J.; HAN, S. Frequency effects on the vowel length contrast merger in seoul korean. *Laboratory Phonology*, v. 6, n. 3-4, p. 469–503, 2015. Disponível em: <<https://doi.org/10.1515/lp-2015-0014>>.
- KAO, D. *Structure of the Syllable in Cantonese*. Mouton, 1971. (Janua linguarum). Disponível em: <<https://books.google.com.br/books?id=0adcuAAACAAJ>>.
- KATZ, J. *Compression effects, perceptual asymmetries, and the grammar of timing*. 2010. Acesso em: 5 set. 2024. Disponível em: <<https://b-linguistics.net>>.
- KAYE, J.; LOWENSTAMM, J.; VERGNAUD, J.-R. Constituent structure and government in phonology. *Phonology*, v. 7, n. 1, p. 193–231, 1990.
- KEATING, P. Phonetic and phonological representation of stop consonant voicing. *Language*, v. 60, 06 1984.
- KELSO, J. A. S.; DEGUZMAN, G. C. The intermittent dynamics of coordination. In: STELMACH, G. E.; REQUIN, J. (Ed.). *Tutorials in motor behavior II*. New York: Elsevier, 1992. p. 549–561.
- KENT, R.; READ, C. *The Acoustic Analysis of Speech*. Singular/Thomson Learning, 2002. ISBN 9780769301129. Disponível em: <<https://books.google.com.br/books?id=OmhqAAAAMAAJ>>.
- KLUENDER, K. R.; DIEHL, R. L.; WRIGHT, B. A. Vowel-length differences before voiced and voiceless consonants: an auditory explanation. *Journal of Phonetics*, v. 16, n. 2, p. 153–169, 1988. ISSN 0095-4470. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0095447019304802>>.
- KUHNERT, B.; HOOLE, P.; MOOSHAMMER, C. Gestural overlap and c-center in selected french consonant clusters. *Proc. 7th International Seminar on Speech Production*, 01 2006.
- KURYLOWICZ, J. Contribution à la théorie de la syllabe. *Biuletyn Polskiego Towarzystwa Językoznawczego (BPTJ)*, v. 8, p. 80–114, 1948.
- LADEFOGED, P.; MADDIESON, I. *The Sounds of the World's Languages*. Oxford: Blackwell Publishers, 1996.

- LAVER, J. *Principles of Phonetics*. [S.l.]: Cambridge University Press, 1994.
- LEHISTE, I. *Consonant quantity and phonological units in Estonian*. Bloomington, IN: Indiana University Press, 1966.
- LEHISTE, I. *Suprasegmentals*. Cambridge, Massachusetts & London, UK: MIT Press, 1970.
- LINDBLOM, B. Phonetic universals in vowel systems. In: OHALA, J.; JAEGER, J. (Ed.). *Experimental Phonology*. [S.l.]: Academic Press, 1986. p. 13–44.
- LISKER, L.; ABRAMSON, A. S. Some effects of context on voice onset time in english stops. *Language and Speech*, v. 10, p. 1–28, 1967.
- LOWENSTAMM, J. Cv as the only syllable type. In: DURAND, J.; LAKS, B. (Ed.). *Current Trends in Phonology: Models and Methods*. [S.l.]: European Studies Research Institute, University of Salford, 1996. p. 419–442.
- MAJOR, R. Stress-timing in brazilian portuguese. *Journal of Phonetics*, v. 9, n. 3, 1981.
- MAJOR, R. Stress and rhythm in brazilian portuguese. *Language*, v. 61, n. 2, p. 259, 1985.
- MAJOR, R. C. Stress and rhythm in brazilian portuguese. In: LOIKE, D. A.; MACEDO, D. P. (Ed.). *Romance Linguistics: The Portuguese Context*. Westport: Bergin and Garvey, 1992. p. 3–30.
- MARCHAL, A.; REIS, C. *Produção da Fala*. 1. ed. Belo Horizonte: UFMG, 2012. v. 1. 297 p.
- MASSINI-CAGLIARI, G. *Acento e Ritmo*. São Paulo: Contexto, 1992.
- MATEUS, M. H. M.; D'ANDRADE, E. *The Phonology of Portuguese*. [S.l.]: Oxford University Press, 2000.
- MCCARTHY, J. J.; PRINCE, A. S. *Prosodic Morphology*. 1986. Manuscript, University of Massachusetts, Amherst I& Brandeis University.
- MEDEIROS, B. R. d. Vogais nasais do português brasileiro: reflexões preliminares de uma revisita. *Revista Letras*, 2007.
- MENDES JÚNIOR, W. A. *A lenição do tepe alveolar intervocálico por falantes do português brasileiro*. Dissertação (Mestrado) — Universidade Federal de Minas Gerais (UFMG), 2018. Acesso Aberto. Disponível em: <<http://hdl.handle.net/1843/LETR-AWUHDA>>.
- MENDONÇA, C. A sílaba em fonologia. *Working Papers em Linguística*, UFSC, v. 7, p. 21–40, 2003.
- MEYER, K. H. *Slavische und indogermanische Intonation*. [S.l.]: Heidelberg, 1920. II. (Sammlung Slavica, II).
- MIGLIORINI, L.; MASSINI-CAGLIARI, G. Sobre o ritmo do português brasileiro: evidências de um padrão acentual. *ReVEL*, v. 8, n. 15, 2010.
- MOL, H.; UHLENBECK, E. The linguistic relevance of intensity in stress. *Lingua*, v. 5, p. 205–213, 1955. ISSN 0024-3841. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0024384155900103>>.

- MORAES, J. A. d. Acentuação lexical e acentuação frasal em português: um estudo acústico perceptivo. In: *Comunicação apresentada no II Encontro Nacional de Fonética e Fonologia*. Brasília: [s.n.], 1986.
- MORAES, J. A. d. Intonation in brazilian portuguese. In: HIRST, D.; CRISTO, A. D. (Ed.). *Intonation systems: a survey of twenty languages*. Cambridge: Cambridge University Press, 1998.
- MOUSIKOU, P.; STRYCHARCZUK, P.; RASTLE, K. Acoustic correlates of stress in speech perception. *Journal of Memory and Language*, v. 136, p. 104509, 2024. ISSN 0749-596X.
- NAM, H.; SALTZMAN, E. A competitive, coupled oscillator model of syllable structure. In: *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona, Spain: [s.n.], 2003. p. 2253–2256.
- NEVADO, H. M. *Relação entre parâmetros acústicos, fonatórios e articulatórios no estabelecimento da proeminência acentual no português brasileiro*. 2020. Apresentação de pôster na Semana da Iniciação Científica - UFMG. (Exposição).
- NOOTEBOOM, S. G. *Production and perception of vowel duration: A study of durational properties of vowels in Dutch*. Tese (Doutorado) — Utrecht University, 1972.
- OHALA, J.; KAWASAKI-FUKUMORI, H. Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In: \_\_\_\_\_. [S.l.: s.n.], 1997. v. 2, p. 343–365. ISBN 9783110805369.
- OKOBI, A. O. A. O. *Acoustic correlates of word stress in American English*. Tese (Doutorado) — Harvard-MIT Division of Health Sciences and Technology, 2006.
- PARKER, S. Sound level protrusions as physical correlates of sonority. *Journal of Phonetics*, v. 36, p. 55–90, 2008.
- PARKER, S. Sonority. In: OOSTENDORP, M. van et al. (Ed.). *The Blackwell Companion to Phonology*. Blackwell Publishing, 2011. Blackwell Reference Online. Accessed 12 January 2012. Disponível em: <[http://www.companiontophonology.com/subscriber/tocnode?id=g9781405184236\\_chunk\\_g978140518423651](http://www.companiontophonology.com/subscriber/tocnode?id=g9781405184236_chunk_g978140518423651)>.
- PASCHEN, L.; FUCHS, S.; SEIFART, F. Final lengthening and vowel length in 25 languages. *Journal of Phonetics*, v. 94, p. 101179, 2022. ISSN 0095-4470. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0095447022000547>>.
- PIKE, K.; PIKE, E. Immediate constituents of mazateco syllables. *International Journal of American Linguistics*, v. 13, p. 78–91, 1947.
- PIKE, K. L. *The Intonation of American English*. Ann Arbor: University of Michigan Press, 1945. v. 1. 34–35 p.
- POLGÁRDI, K. The representation of sc-clusters: A strict cv analysis. *Glossa: a journal of general linguistics*, v. 7, n. 1, p. 1–29, 2022. Disponível em: <<https://doi.org/10.16995/glossa.5702>>.

- PRADO, V. A. M.; CANTONI, M. M. Stress-induced effects in coda duration: an analysis of brazilian portuguese through articulatory phonology. In: JAURENA, M.; METZLER, Z. (Ed.). *Proceedings of the 36th Western Conference on Linguistics*. Fresno, CA: California State University, Fresno, 2024. (WECOL, v. 30), p. 1. Hosted by Department of Linguistics, California State University, Fresno.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>.
- ROESSIG, S.; MÜCKE, D. Modeling dimensions of prosodic prominence. *Frontiers in Communication*, v. 4, 2019. ISSN 2297-900X. Disponível em: <<https://www.frontiersin.org/journals/communication/articles/10.3389/fcomm.2019.00044>>.
- RONQUEST, R. E. An acoustic examination of unstressed vowel reduction in heritage spanish. In: HOWE, C. et al. (Ed.). *Selected Proceedings of the 15th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, 2013. p. 157–171. Document #2882, available at <[www.lingref.com](http://www.lingref.com)>.
- ROUCH, M.; LUNDEN, A. The status of word-final phonetic phenomena. *Proceedings of the Linguistic Society of America*, v. 5, n. 1, p. 599–606, Mar 2020. Disponível em: <<https://journals.linguisticsociety.org/proceedings/index.php/PLSA/article/view/4730>>.
- SAHKAI, H.; MIHKLA, M. Acoustic correlates of primary word stress in estonian. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, v. 15, n. 2, p. 5–30, jun. 2024. Disponível em: <<https://ojs.utlib.ee/index.php/jeful/article/view/23790>>.
- SALTZMAN, E.; BYRD, D. Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, v. 19, n. 4, p. 499–526, 2000. ISSN 0167-9457. Debates in Dynamics. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167945700000300>>.
- SALTZMAN, E.; KELSO, S. Skilled actions: A task-dynamic approach. *Psychological review*, v. 94, p. 84–106, 01 1987.
- SARAN, F. *Deutsche verslehre*. C.H. Beck'sche verlagsbuchhandlung. O. Beck, 1907. (Deutsche verslehre, v. 3, pt. 3). Disponível em: <<https://books.google.com.br/books?id=bPkqAAAAMAAJ>>.
- SCHMEISER, B. On the current state of vowel intrusion analysis in spanish within optimality theory. In: AL., C. et (Ed.). *Selected Proceedings of the 11th Hispanic Linguistics Symposium*. Somerville, MA: Cascadilla Proceedings Project, 2009. p. 207–218.
- SCHMITT, A. *Untersuchungen zur allgemeinen Akzentlehre*. Heidelberg: Winter, 1924.
- SCHOOOR, A.; AICHERT, I.; ZIEGLER, W. A motor learning perspective on phonetic syllable kinships: How training effects transfer from learned to new syllables in severe apraxia of speech. *Aphasiology*, in press, p. –, 07 2012.
- SELKIRK, E. The syllable. In: HULST, H. van der; SMITH, N. (Ed.). *The Structure of Phonological Representations (Part II)*. Dordrecht: Foris, 1982. p. 337–383.
- SELKIRK, E. On the major class features and syllable theory. In: ARONOFF, M.; OEHRLE, R. T. (Ed.). *Language and Sound Structure*. Cambridge, MA: MIT Press, 1984. p. 107–136.

- SEVERIJNEN, G. G.; BOSKER, H. R.; MCQUEEN, J. M. Your “voornaam” is not my “voornaam”: An acoustic analysis of individual talker differences in word stress in dutch. *Journal of Phonetics*, v. 103, p. 101296, 2024. ISSN 0095-4470. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0095447024000020>>.
- SILVA, T. C. *Fonética e fonologia do português: Roteiro de estudos e guia de exercícios*. [S.l.]: Contexto, 2003.
- SILVEIRA, F.; SEARA, I. C. Vogal de apoio em grupos consonantais ccv no português brasileiro. *Revista da ABRALIN*, v. 7, n. 1, maio 2017. Disponível em: <<https://revista.abralin.org/index.php/abralin/article/view/977>>.
- SLUIJTER, A.; HEUVEN, V. V. Spectral tilt as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, v. 100, p. 2471–85, 11 1996.
- SLUIJTER, A. M. C.; HEUVEN, V. J. van. Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in dutch. *Phonetica*, v. 52, p. 71–89, 1995.
- SOUSA, E. M. G. d. *Para a caracterização fonético-acústica da nasalidade no português do Brasil*. Dissertação (Dissertação (Mestrado)) — Universidade Estadual de Campinas, 1994.
- STERIADE, D. *Positional neutralization and the expression of contrast*. 1994. Manuscript, University of California, Los Angeles; revised text of 1993 NELS talk.
- STETSON, R. H. *Motor Phonetics: A Study of Speech Movements in Action*. 2. ed. Amsterdam: North Holland Publishing Co., 1951.
- TEVDORADZE, I. *Kartuli enis p'rosodiis sak'itxebi [Prosodic issues in the Georgian language]*. Tbilisi: Tbilisi State University Publishing, 1978.
- THIESSEN, E. D.; SAFFRAN, J. R. When cues collide: use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology*, v. 39, n. 4, p. 706–716, 2003. Disponível em: <<https://doi.org/10.1037/0012-1649.39.4.706>>.
- TILSEN, S. Space and time in models of speech rhythm. *Annals of the New York Academy of Sciences*, Wiley, 2019.
- TILSEN, S. et al. A cross-linguistic investigation of articulatory coordination in word-initial consonant clusters. *Cornell Working Papers in Phonetics and Phonology*, v. 2012, p. 51–81, 2012.
- TOPINTZI, N. *Onsets: suprasegmental and prosodic behaviour*. [S.l.]: Cambridge University Press, 2010. (Cambridge Studies in Linguistics, 125).
- TOPINTZI, N. Onsets. In: OOSTENDORP, M. et al. (Ed.). *The Blackwell Companion to Phonology*. [S.l.]: Blackwell, 2011.
- TSCHENKELI, K. *Einführung in die georgische Sprache*. Zürich: Amirani Verlag, 1958.
- VOGEL, I.; ATHANASOPOULOU, A.; PINKUS, N. Prominence, contrast, and the functional load hypothesis: An acoustic investigation. In: HEINZ, J.; GOEDEMANS, R.; HULST, H. van der (Ed.). *Dimensions of Phonological Stress*. Cambridge: Cambridge University Press, 2016. p. 123–167.

- WARNER, N. et al. Methods for studying spontaneous speech. In: \_\_\_\_\_. [S.l.]: Oxford University Press, 2012. p. 612–633.
- WETZELS, W. L. Mid vowel neutralization in brazilian portuguese. *Cadernos de Estudos Linguísticos*, v. 23, p. 19–55, 1992.
- WETZELS, W. L. The lexical representation of nasality in brazilian portuguese. *Probus*, v. 92, 1997.
- WETZELS, W. L. The representation of vowel height and vowel height neutralization in brazilian portuguese (southern dialects). *Tones and Features*, p. 331–360, 2011.
- WILLIAMS, B. The phonetic manifestation of stress in welsh. In: HULST, H. G. van der (Ed.). *Word Prosodic Systems in the Languages of Europe*. Berlin: Mouton de Gruyter, 1999. p. 311–334.
- XU, Y. In defense of lab speech. *Journal of Phonetics*, v. 38, n. 3, p. 329–336, 2010. ISSN 0095-4470. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0095447010000318>>.
- YOUNG, S. et al. The htk book. *Cambridge University Engineering Department*, v. 3, 2002.
- ZHGHENTI, S. Kartuli dialekt'ebis c'armotkmis taviseburebata šesc'avlis zogierti sak'itxi [some questions about the study of pronounciational properties of georgian dialects]. *St'alinis saxelobis Tbilisis Saxelmts'ipo Universit'et'is šromebi*, v. 69, p. 253–264, 1958.

# Apêndices

# APÊNDICE A – Parecer consubstanciado - Projeto aprovado

## PARECER CONSUBSTANCIADO DO CEP

### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** O acento lexical e seu impacto nos constituintes silábicos

**Pesquisador:** MARIA MENDES CANTONI

**Área Temática:**

**Versão:** 1

**CAAE:** 80099224.6.0000.5149

**Instituição Proponente:** UNIVERSIDADE FEDERAL DE MINAS GERAIS

**Patrocinador Principal:** Financiamento Próprio

### DADOS DO PARECER

**Número do Parecer:** 6.908.402

#### **Apresentação do Projeto:**

O estudo investiga os efeitos acústicos do acento de palavra no português brasileiro (PB) em diferentes moldes silábicos. Trabalhos que se propõem a estudar o acento de palavra nas línguas são limitados à análise do molde silábico mais natural e esperado: o CV. Essa delimitação parece partir da desconsideração do papel de segmentos não-vocálicos em contribuírem com a manifestação acústica do acento de palavra. Um agravante é que, dentro da literatura referente ao PB, os trabalhos dão prioridade exclusiva à análise fonológica, à qual escapa medidas acústicas do fenômeno, como a duração. A hipótese é que a duração dos segmentos acentuados será afetada pelo tipo de molde silábico, com sílabas mais simples mostrando um aumento mais latente de duração. Além disso, esperamos observar variações na distribuição desse aumento de duração entre os diferentes moldes silábicos, bem como um papel não descartável das consoantes na manifestação acústica do acento de palavra em PB.

#### **Objetivo da Pesquisa:**

Objetivo Primário:

Avaliar o efeito do molde silábico nas marcas acústicas do acento.

Objetivo Secundário:

Contribuir com a avaliação do acento de palavra em PB, propondo uma abordagem inovadora

**Endereço:** Av. Presidente Antonio Carlos, 6627 2º. Andar Sala 2005 Campus Pampulha

**Bairro:** Unidade Administrativa II **CEP:** 31.270-901

**UF:** MG **Município:** BELO HORIZONTE

**Telefone:** (31)3409-4592

**E-mail:** coep@prpq.ufmg.br

Continuação do Parecer: 6.908.402

que integra a análise detalhada dos constituintes silábicos, vocálicos e não-vocálicos, em diferentes moldes silábicos.

**Avaliação dos Riscos e Benefícios:**

Riscos:

Durante a participação no estudo, é possível que o sujeito experimente sensações de cansaço ou constrangimento devido ao monitoramento. Se isso ocorrer, ele(a) terá a opção de fazer uma pausa ou interromper a participação a qualquer momento, sem prejudicar o andamento do estudo ou alterar a relação com o pesquisador responsável.

Benefícios:

É importante destacar que a participação no estudo não resultará em benefícios diretos pessoais, entretanto, contribuirá significativamente para a compreensão dos sons presentes no português falado no Brasil. A decisão de participar é voluntária, e não será fornecida nenhuma remuneração ou cobrança associada. Em caso de qualquer dano decorrente da participação no estudo, o participante terá o direito de buscar indenização, assegurando assim seus direitos e bem-estar.

**Comentários e Considerações sobre a Pesquisa:**

O projeto é bom e claro, seu mérito foi julgado e aprovado pela unidade acadêmica do proponente. Quanto aos aspectos éticos em pesquisa com seres humanos, o projeto descreve seus procedimentos, formas de mitigação de possíveis danos, assim como instrumentos de coleta compatíveis.

**Considerações sobre os Termos de apresentação obrigatória:**

Foram apresentados:

- Formulário de informações básicas;
- Folha de rosto assinada;
- Parecer de aprovação na unidade acadêmica;
- Projeto completo;
- TCLE;
- Descrição dos instrumentos de coleta.

**Recomendações:**

No TCLE o pesquisador assistente está como responsável/coordenador, nesse caso, alterar a denominação ou o nome do pesquisador que irá assinar o documento.

**Endereço:** Av. Presidente Antonio Carlos, 6627 2º. Andar Sala 2005 Campus Pampulha

**Bairro:** Unidade Administrativa II **CEP:** 31.270-901

**UF:** MG **Município:** BELO HORIZONTE

**Telefone:** (31)3409-4592

**E-mail:** coep@prpq.ufmg.br

Continuação do Parecer: 6.908.402

Ainda no TCLE, incluir campos para assinatura/rubrica em todas as páginas.

**Conclusões ou Pendências e Lista de Inadequações:**

Reconenda-se aprovação do projeto.

**Considerações Finais a critério do CEP:**

Tendo em vista a legislação vigente (Resolução CNS 466/12), o CEP-UFMG recomenda aos Pesquisadores: comunicar toda e qualquer alteração do projeto e do termo de consentimento via emenda na Plataforma Brasil, informar imediatamente qualquer evento adverso ocorrido durante o desenvolvimento da pesquisa (via documental encaminhada em papel), apresentar na forma de notificação relatórios parciais do andamento do mesmo a cada 06 (seis) meses e ao término da pesquisa encaminhar a este Comitê um sumário dos resultados do projeto (relatório final).

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_2342837.pdf	24/05/2024 20:04:42		Aceito
Parecer Anterior	Parecer_Consubstanciado_assinado.pdf	24/05/2024 20:04:18	VITOR ANTONIO MORAIS PRADO	Aceito
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_2342837.pdf	24/05/2024 14:47:51		Aceito
Projeto Detalhado / Brochura Investigador	ProjetoMestrado_VitorPrado.pdf	24/05/2024 14:45:32	VITOR ANTONIO MORAIS PRADO	Aceito
Projeto Detalhado / Brochura Investigador	ProjetoMestrado_VitorPrado.pdf	24/05/2024 14:45:32	VITOR ANTONIO MORAIS PRADO	Postado
Outros	Instrumento_de_coleta.pdf	24/05/2024 14:42:27	VITOR ANTONIO MORAIS PRADO	Aceito
Outros	Instrumento_de_coleta.pdf	24/05/2024 14:42:27	VITOR ANTONIO MORAIS PRADO	Postado
Folha de Rosto	folha_de_rosto_assinada.pdf	24/05/2024 14:04:23	VITOR ANTONIO MORAIS PRADO	Aceito
Folha de Rosto	folha_de_rosto_assinada.pdf	24/05/2024 14:04:23	VITOR ANTONIO MORAIS PRADO	Postado
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_header.pdf	23/05/2024 08:32:33	VITOR ANTONIO MORAIS PRADO	Aceito

**Endereço:** Av. Presidente Antonio Carlos, 6627 2º. Andar Sala 2005 Campus Pampulha

**Bairro:** Unidade Administrativa II

**CEP:** 31.270-901

**UF:** MG

**Município:** BELO HORIZONTE

**Telefone:** (31)3409-4592

**E-mail:** coep@prpq.ufmg.br

Continuação do Parecer: 6.908.402

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

BELO HORIZONTE, 25 de Junho de 2024

---

**Assinado por:**  
**Corinne Davis Rodrigues**  
**(Coordenador(a))**

**Endereço:** Av. Presidente Antonio Carlos, 6627 2º. Andar Sala 2005 Campus Pampulha

**Bairro:** Unidade Administrativa II **CEP:** 31.270-901

**UF:** MG **Município:** BELO HORIZONTE

**Telefone:** (31)3409-4592

**E-mail:** coep@prpq.ufmg.br

# Anexos

# ANEXO A – Termo de Consentimento Livre e Esclarecido (TCLE)

# **Termo de Consentimento Livre e Esclarecido (TCLE)**

## **Descrição do Estudo**

Você está sendo convidado(a) a participar da pesquisa “O acento lexical e seu impacto nos constituintes silábicos”, coordenada por Maria Mendes Cantoni. O objetivo deste estudo é avaliar os efeitos do acento tônico na duração de vogais e de consoantes no Português Brasileiro.

## **Procedimentos**

Sua participação consistirá em visualizar uma série de sentenças, que aparecerão na tela do computador à sua frente. Após ler cada frase, pediremos que você a repita com entoação de uma frase normal, sem incidir foco sobre alguma das palavras.

Registraremos suas respostas por áudio no computador e depois serão feitas avaliações para caracterizar como os sons do português são produzidos.

A coleta de dados será realizada na Faculdade de Letras da UFMG. A duração da sessão de coleta será de 20 minutos, aproximadamente. Caso você tenha que se deslocar ao local da coleta especificamente para a sua participação, será ressarcido o valor gasto com deslocamento.

## **Riscos e Benefícios**

Durante sua participação no estudo, é possível que você experimente sensações de cansaço ou constrangimento devido ao monitoramento. Se isso ocorrer, você terá a opção de fazer uma pausa ou interromper sua participação a qualquer momento, sem prejudicar o andamento do estudo ou alterar a relação com o pesquisador responsável.

É importante destacar que sua participação não resultará em benefícios diretos pessoais, entretanto, contribuirá significativamente para a compreensão dos sons presentes no português falado no Brasil. A decisão de participar é voluntária, e não será fornecida nenhuma remuneração ou cobrança associada. Em caso de qualquer dano decorrente de sua participação no estudo, você terá o direito de buscar indenização, assegurando assim seus direitos e bem-estar.

---

Assinatura do Participante

## **Confidencialidade e Armazenamento de Dados**

Os dados adquiridos durante o estudo serão mantidos pelo pesquisador responsável em seu computador na Faculdade de Letras, sendo retidos pelo período de 5 a 10 anos. Não haverá compartilhamento desses dados com terceiros, exceto para os outros pesquisadores envolvidos na análise do projeto, e isso será feito exclusivamente de maneira anonimizada. Essas medidas visam garantir a confidencialidade e a segurança dos dados coletados, protegendo a privacidade dos participantes da pesquisa.

## **Contato**

Para qualquer dúvida em relação ao estudo, o pesquisador responsável, Vitor Antônio M. Prado, pode ser contactado pelo e-mail [vitormoraisprado@gmail.com](mailto:vitormoraisprado@gmail.com) ou pelo telefone celular (31) 98606-0125.

Para dúvidas relacionadas a aspectos éticos da participação no estudo, pode ser contactado o COEP UFMG, no endereço AV. Presidente Antônio Carlos, 6627, Pampulha - Belo Horizonte - MG - CEP 31270-901. Unidade Administrativa I | - 2º Andar - Sala: 2005. Telefone: (031) 3409-4592. Horário de atendimento: 09:00 às 11:00 / 14:00 às 16:00.

## **AO PARTICIPANTE**

Declaro que minhas dúvidas foram esclarecidas pelo pesquisador e que concordo em participar de forma voluntária deste estudo

Dados para Contato do Participante:

Nome: \_\_\_\_\_.

Endereço: \_\_\_\_\_.

E-mail: \_\_\_\_\_.

Telefone(s): (\_\_\_\_) \_\_\_\_\_-\_\_\_\_ / (\_\_\_\_) \_\_\_\_\_-\_\_\_\_.

---

Assinatura do Pesquisador

Maria Mendes Cantoni

---

Assinatura do Participante

# ANEXO B – Conjunto de dados completo

NaN0>0 1NaN0>0 1NaN0>0 1

Tipo	Modo	Ponto	Vozeamento Molde sí- lábico	Palavra- veículo (átona)	Palavra- veículo (tônica)
------	------	-------	-----------------------------------	-----------------------------	---------------------------------

(Continued)

|

capa

Amapá

(Continued)

|

rosas

rosás

(Continued)

Consoante Oclusiva Velar	Vozeada	CVC CV+s	g	cegas	ingás
--------------------------	---------	-------------	---	-------	-------

(Continued)

|

ladra

empedrá

(Continued)

Consoante	Oclusiva Alveolar	Desvozeada	CCVC	t	mantras	entrás
			C+r+V+s			

# ANEXO C – Script em R para geração de apresentações

```
1 # Instala e carrega pacotes necessarios
2 #install.packages("xml2")
3 #install.packages('officer')
4 #install.packages('magrittr')
5 library(officer)
6 library(magrittr)
7 library(xml2)
8
9 # O script a seguir ira gerar apresentacoes diferentes, com slides de
  teste ordenados aleatoriamente
10 # Para mudar o numero de apresentacoes, mude o numero n_apres a
  seguir:
11 n_apres <- 20
12 for (k in 1:n_apres){
13   fileout <- tempfile(fileext = ".pptx")
14   doc <- read_pptx()
15
16   # Especifique o caminho para a pasta com os arquivos contendo as
  frases de teste e treinamento
17   setwd("C:/Users/Vitor Prado/Desktop/Mestrado/piloto/data_experiment
  ")
18   set.seed(k)
19
20   treino <- readLines("frases_treino.txt", encoding="UTF-8")
21
22   # Adiciona slides com frases de treinamento
23   n_treino <- 3
24   for (i in 1:n_treino){
25     doc <- add_slide(doc)
26     doc <- ph_with(x = doc, value = treino[i], location = ph_location
  (1,3,8,1))
27   }
28
29   # Adiciona um slide em branco (opcional)
30   doc <- add_slide(doc, layout = "Title Only", master = "Office Theme
  ")
```

```
31
32 # Adiciona slides com frases de teste
33 n_teste <- 264
34 dados <- readLines("frases_randomizadas.txt", encoding="UTF-8")
35 set.seed(k)
36 dados <- sample(dados, n_teste)
37 write.table(dados, file = paste("ApresentacaoVitor_", k, ".txt",
38     sep=""), append = FALSE, sep = ",", dec = ".",
39     row.names = FALSE, col.names = FALSE)
40 for (i in 1:n_teste){
41     doc <- add_slide(doc)
42     doc <- ph_with(x = doc, value = dados[i], location = ph_location
43         (1,3,8,1))
44 }
45 # Salva o arquivo PowerPoint gerado
46 print(doc, target = paste("ApresentacaoVitor_", k, ".pptx", sep="")
47     )
48 }
```

## ANEXO D – Script em Praat para remoção de semivogal palatal

```
1 #Retirar [j] das palavras
2
3 #Selecione o TextGrid, um por vez
4
5 Replace interval texts: 1, 1, 0, "aj", "a", "Regular Expressions"
6 Replace interval texts: 1, 1, 0, "baSi~Ju", "bajSi~Ju", "Regular
  Expressions"
```

## ANEXO E – Script em Praat para organizar camada de palavras do TextGrid

```
1 ###  
2  
3 #Selecione o TextGrid, um por vez  
4  
5 numtier = Get number of tiers  
6  
7 #Duplica e limpa o tier de palavras  
8 Duplicate tier: 3, 4, "mainwords"  
9 Replace interval texts: 4, 1, 0, "eu|digo|baixinho|_", "", "Regular  
   Expressions"  
10  
11 ###
```

## ANEXO F – Script em Praat para marcar o vozeamento de [s]

```

1 ###
2
3 writeInfoLine: "Extracting..."
4
5 directory$ = "C:/Users/Vitor Prado/Desktop/Mestrado/piloto/data_
   experiment/ udios cortados + textgrids"
6 directory$ = "'directory$'" + "/"
7
8 # Create a list of .wav files in the selected folder
9 Create Strings as file list: "wavList", directory$ + "*.wav"
10
11 # Get the number of .wav files found
12 numberOfWavs = Get number of strings
13
14 # Loop through each .wav file
15 for i from 1 to numberOfWavs
16     # Get the current .wav file name
17     Create Strings as file list: "wavList", directory$ + "*.wav"
18     wavFile$ = Get string: i
19     clearinfo
20     appendInfoLine: "Processing ", wavFile$
21
22     # Construct the corresponding .TextGrid file name (assumes the
       same base name as .wav)
23     textGridFile$ = replace$ (wavFile$, ".wav", ".TextGrid", 0)
24
25     # Open the .wav and .TextGrid files
26     Read from file: directory$ + wavFile$
27     thisSound$ = selected$("Sound")
28
29     Read from file: directory$ + textGridFile$
30     thisTextGrid$ = selected$("TextGrid")
31
32
33     # Extract the number of intervals in the phoneme tier
34     select TextGrid 'thisTextGrid$'

```

```
35     numberOfPhonemes = Get number of intervals: 1
36
37     # Insert a new interval tier for voicing
38     Insert interval tier: 7, "voicing"
39     voicingtier = 7
40
41     # Create the [____] object for the current sound file
42     select Sound 'thisSound$'
43     #To Pitch: 0.01, 75, 600
44     To Formant (burg): 0, 5, 5500, 0.015, 50
45
46
47     # Loop through each interval on the phoneme tier
48     for thisInterval from 1 to numberOfPhonemes
49         # Get the label of the interval
50         select TextGrid 'thisTextGrid$'
51         thisPhoneme$ = Get label of interval: 1, thisInterval
52
53         nextInterval = thisInterval + 1
54
55         if nextInterval < numberOfPhonemes
56             nextPhoneme$ = Get label of interval: 1, nextInterval
57
58             if thisPhoneme$ = "s" and nextPhoneme$ = "b"
59                 # Find the start, midpoint, and duration of the
60                 interval
61                 thisPhonemeStartTime = Get start point: 1,
62                 thisInterval
63                 thisPhonemeEndTime = Get end point: 1, thisInterval
64                 duration = thisPhonemeEndTime - thisPhonemeStartTime
65
66                 # Calculate the 2/4 and 3/4 points of the interval
67                 twoFourPoint = thisPhonemeStartTime + duration * 1 /
68                 4
69                 threeFourPoint = thisPhonemeStartTime + duration * 3
70                 / 4
71
72                 # Get the mean pitch between 2/4 and 3/4 of the
73                 interval
74                 #select Pitch 'thisSound$'
75                 #meanPitch = Get value at time: threeFourPoint, "
76                 Hertz", "linear"
```

```
71
72     select Formant 'thisSound$'
73     meanF1 = Get mean: 1, twoFourPoint,
74             thisPhonemeEndTime, "hertz"
75     meanF1$ = string$ (meanF1)
76
77
78     appendInfoLine: thisInterval, thisPhoneme$, meanF1
79
80     #if meanPitch > 0
81     if meanF1 < 600
82         select TextGrid 'thisTextGrid$'
83         Set interval text: 1, thisInterval, "z"
84
85         # Optionally insert boundaries and label in the
86         # new voicing tier
87         Insert boundary: voicingtier,
88             thisPhonemeStartTime
89         Insert boundary: voicingtier, thisPhonemeEndTime
90         voicedint = Get interval at time: voicingtier,
91             thisPhonemeStartTime
92         Set interval text: voicingtier, voicedint, meanF1
93         $
94     endif
95     endif
96     endif
97     endfor
98
99     # Save the modified TextGrid file
100    select TextGrid 'thisTextGrid$'
101    Save as text file: directory$ + "textgrids processados/" +
102        textGridFile$
103
104    endfor
105
106    clearinfo
107    appendInfoLine: newline$, newline$, "Whoo-hoo! Processing complete!"
108
109    ###
```

## ANEXO G – Script em Praat para extrair a duração de intervalos dos TextGrids

```

1 file_type$ = ".TextGrid"
2 directory$ = "C:/Users/Vitor Prado/Desktop/Mestrado/piloto/data_
   experiment/ udios cortados + textgrids"
3 directory$ = "'directory$'" + "/"
4 resultfile$ = "'directory$'" + "all.txt"
5
6 #Read from file... E:/Luciana_Vitor/editados/S014M.TextGrid
7 #sn$ = selected$ ("TextGrid")
8 #selectObject: "TextGrid 'sn$'"
9
10 filedelete 'resultfile$'
11
12 clearinfo
13
14 header_row$ = "part" + tab$ + "ortholabel" + tab$ + "seglabel" + tab$
   + "sylllabel" + tab$ + "wordlabel" + tab$ + "segdur" + tab$ + "
   sylldur" + tab$ + "worddur" + newline$
15
16 #header_row$ = "part" + tab$ + "label" + tab$ + "startint" + tab$ + "
   endint" + tab$ + "intduration" + newline$
17
18 fileappend "'resultfile$'" 'header_row$'
19
20 #Camada de referencia:
21 orthotier = 6
22 wordtier = 4
23 sylltier = 2
24 segtier = 1
25
26 Create Strings as file list... list 'directory$'*'.TextGrid
27 n_textgrids = Get number of strings
28
29 if !n_textgrids
30     exit There are no .TextGrid files in the folder!
31 endif
32

```

```
33 for ifile from 1 to n_textgrids
34     select Strings list
35     curr_file$ = Get string... 'ifile'
36     Read from file... 'directory$' 'curr_file$'
37
38     # Here we make a variable called "object_name$" that will be
39     # equal to the filename minus the ".wav" extension
40     sn$ = selected$ ("TextGrid")
41
42     numortho = Get number of intervals: orthotier
43     numint = Get number of intervals: wordtier
44     numint2 = numint - 1
45
46     writeInfoLine: "Reading file 'curr_file$': 'ifile'/'n_
47     textgrids'"
48
49     for i from 1 to numint2
50         #appendInfoLine: "for = ", "'i'"
51         wordlabel$ = Get label of interval: wordtier, i
52
53         if wordlabel$ <> "" and wordlabel$ <> "_"
54             wordstartint = Get start time of interval:
55             wordtier, i
56             wordendint = Get end time of interval:
57             wordtier, i
58             worddur = wordendint - wordstartint
59             firstsyll = Get high interval at time:
60             sylltier, wordstartint
61             sylllabel$ = Get label of interval: sylltier,
62             firstsyll
63             syllstartint = wordstartint
64             syllendint = Get end time of interval:
65             sylltier, firstsyll
66
67             nsyll = firstsyll
68
69             sylldur = syllendint - syllstartint
70
71             numortho = Get interval at time: orthotier,
72             wordstartint
73             ortholabel$ = Get label of interval:
74             orthotier, numortho
```

```
66
67     while syllendint <= wordendint
68     #appendInfoLine: "primeiro while = ", "'nsyll
69         ' "
70
71         firstseg = Get high interval at time:
72             segtier, syllstartint
73         seglabel$ = Get label of interval:
74             segtier, firstseg
75         segstartint = syllstartint
76         segendint = Get end time of interval:
77             segtier, firstseg
78
79         nseg = firstseg
80
81         segdur = segendint - segstartint
82
83         #appendInfoLine: "segendint = ", "'
84             segendint' "
85         #appendInfoLine: "syllendint = ", "'
86             syllendint' "
87
88     while segendint <= syllendint
89         #appendInfoLine: "segundo
90             while = ", "'nseg' "
91
92         #Colar informa es
93         outputLine$ = "'sn$' " + tab$
94             + "'ortholabel$' " + tab$ + "'
95             + "'seglabel$' " + tab$ + "'
96             + "'sylllabel$' " + tab$ + "'
97             + "'wordlabel$' " + tab$ + "'
98             + "'segdur' " + tab$ + "'
99             + "'sylldur' " + tab$ + "'
100             + "'worddur' "
101         appendFileLine: resultfile$,
102             outputLine$
103
104         #Incrementar o valor da label
105         nseg = nseg + 1
```

```
92         seglabel$ = Get label of
93             interval: segtier, nseg
94         segstartint = Get start time
95             of interval: segtier, nseg
96         segendint = Get end time of
97             interval: segtier, nseg
98         segdur = segendint -
99             segstartint
100
101         #Incrementar o valor da label
102         nsyll = nsyll + 1
103         sylllabel$ = Get label of interval:
104             sylltier, nsyll
105         syllstartint = Get start time of
106             interval: sylltier, nsyll
107         syllendint = Get end time of interval
108             : sylltier, nsyll
109
110         sylldur = syllendint - syllstartint
111
112     endwhile
113 endif
114 endfor
115 endlfor
116 printline Script concluído.
```