

# FILOLOGIA E LINGUÍSTICA PORTUGUESA

26 (2) ago./dez., 2024

Universidade de São Paulo  
Faculdade de Filosofia, Letras e Ciências Humanas  
Pós-graduação em Filologia e Língua Portuguesa

e-ISSN 2176-9419

FILOLOGIA E  
LINGUÍSTICA  
PORTUGUESA

26(2), ago./dez. 2024  
Estudos em Fonética e Fonologia

Universidade de São Paulo  
Faculdade de Filosofia, Letras e Ciências Humanas  
Pós-graduação em Filologia e Língua Portuguesa

e-ISSN: 2176-9419

# Revista Filologia e Linguística Portuguesa

Universidade de São Paulo  
Faculdade de Filosofia, Letras e Ciências Humanas  
Pós-graduação em Filologia e Língua Portuguesa



## **Editores**

Sílvio de Almeida Toledo Neto, *Universidade de São Paulo*, Brasil

Maria Clara Paixão de Sousa, *Universidade de São Paulo*, Brasil

Paulo Ângelo Araújo-Adriano, *Universidade de São Paulo*, Brasil

## **Conselho Editorial**

Ana Rosa Ferreira Dias, *Pontifícia Universidade Católica de São Paulo*, Brasil

Anthony Julius Naro, *Universidade Federal do Rio de Janeiro*, Brasil

Bertrand Daunay, *Université de Lille*, França

Cristina Sobral, *Universidade de Lisboa*, Portugal

Esperança Carneira, *Universidade de Lisboa*, Portugal

Evanildo Bechara, *Universidade do Estado do Rio de Janeiro*, Brasil

Fabiana Cristina Komesu, *Universidade Estadual Paulista Júlio de Mesquita Filho*, Brasil

Françoise Boch, *Université de Grenoble Les Alpes*, França

Graça Maria Rio-Torto, *Universidade de Coimbra*, Portugal

João Wanderley Geraldi, *Universidade de Campinas*, Brasil

Juliana Alves Assis, *Pontifícia Universidade Católica de Minas Gerais*, Brasil

Leda Bisol, *Universidade Federal do Rio Grande do Sul*, Brasil

Leonor Lopes Fávero, *Pontifícia Universidade Católica de São Paulo*, Brasil

Luciani Ester Tenani, *Universidade Estadual Paulista Júlio de Mesquita Filho*, Brasil

Maria Filomena Candeias Gonçalves, *Universidade de Évora*, Portugal

Maria Helena de Moura Neves, *Universidade Estadual Paulista Júlio de Mesquita Filho*, Brasil

Maria Teresa Lino, *Universidade Nova de Lisboa*, Portugal

Marina Vigário, *Universidade de Lisboa*, Portugal

Marisa Cruz, *Universidade de Lisboa*, Portugal

Mary Kato, *Universidade de Campinas*, Brasil

Raquel Salek Fiad, *Universidade de Campinas*, Brasil

Rodolfo Ilari, *Universidade de Campinas*, Brasil

Sónia Frota, *Universidade de Lisboa*, Portugal

Ataliba Teixeira de Castilho, *Universidade de São Paulo*, Brasil

Ieda Maria Alves, *Universidade de São Paulo*, Brasil

Manoel Luiz Gonçalves Correa, *Universidade de São Paulo*, Brasil

Maria Cristina Altman, *Universidade de São Paulo*, Brasil

A Revista *Filologia e Linguística Portuguesa* publica estudos em Filologia e Linguística, com atenção ao seu valor para as investigações sobre a linguagem em geral e a língua portuguesa em particular. Os trabalhos cobrem linhas diversas, como a crítica textual, a paleografia, a codicologia, a linguística histórica, os estudos gramaticais, a análise do discurso e a historiografia linguística.

Editada desde 1997, a Revista é apoiada pelo Programa de Pós-Graduação em Filologia e Língua Portuguesa do Departamento de Letras Clássicas e Vernáculas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo.

A revista é uma publicação exclusivamente digital, com acesso exclusivo pelo Portal de Periódicos da Universidade de São Paulo, em <http://www.revistas.usp.br/flp>.

e-ISSN: 2176-9419

Copyright (c) 2024 Filologia e Linguística Portuguesa



Esta obra possui uma licença *Creative Commons Attribution-NonCommercial 4.0 International License*. Os direitos autorais serão cedidos à revista para publicação on-line com livre acesso e impressa para arquivo em papel. Serão preservados, porém, para autores que queiram republicar seus trabalhos em coletâneas.

Filologia e Linguística Portuguesa,  
volume 26, número 2  
Estudos em Fonética e Fonologia

*Editores convidados*

Ronaldo Manguiera Lima Júnior  
Flaviane Romani Fernandes Svartman

Sumário

*Apresentação*  
**179-182**

*Reestruturação de encontros obstruinte-líquida*  
Maria Cantoni  
**183-198**

*O alçamento não assimilatório das vogais /e, o/ pretônicas no português do sul do Brasil: uma abordagem fonológica de contraste e aperfeiçoamento*  
Elisa Battisti  
**199-219**

*Os efeitos do aprendizado das letras <e, o> na produção oral de vogais médias pretônicas e postônicas finais*  
Cecília Toledo  
**221-239**

*O processamento lexical do acento em pseudopalavras a partir da leitura de falantes do português brasileiro*  
Aline Benevides  
**241-259**

*Sândi vocálico externo em português brasileiro: elisão, ditongação e coalescência*

Magnun Madruga e Guilherme Gonçalves

**261-284**

*Vogais tônicas orais do kabuverdianu do Príncipe: uma análise fonética*

Shirley Freitas e Amanda Balduino

**285-308**

*O onset no angolar moderno*

Manuele Bandeira e Ubiratã Alves

**309-327**

*Correlações entre percepção e produção de palavras heterotônicas do espanhol por aprendizes brasileiros*

Pollianna Millan

**329-348**

*Desenvolvimento dos sons vocálicos do inglês como L2 por brasileiros em contexto de imersão*

Anderson Romário Souza Silva, Ronaldo Mangueira Lima Júnior, Lucas Heitor Ananias Oliveira e Dominyk Santos Dias Simões

**349-374**

*Diferenças na produção de foco prosódico contrastivo na fala de adultos e de crianças com aquisição fonológica típica e atípica do Português Brasileiro*

Geovana Soncin, Cecília Lorena Silva Guida, Fernanda Leitão de Castro Nunes de Lima e Larissa Cristina Berti

**375-395**

*Os impactos vocais da doença de Parkinson em idosos*

Lucas Manca Dal'Ava e Plínio Almeida Barbosa

**397-418**

## Apresentação

### Estudos em Fonética e Fonologia

O volume 26, número 2, da revista *Filologia e Linguística Portuguesa* apresenta um conjunto de onze trabalhos que abordam temas concernentes à área de Fonética e Fonologia. Trata-se de um volume temático que abrange, em sua grande maioria, artigos que são frutos de trabalhos apresentados no V Encontro Intermediário do Grupo de Trabalho (GT) de Fonética e Fonologia da Associação Nacional de Pós-graduação e Pesquisa em Letras e Linguística (ANPOLL), realizado na Universidade de São Paulo (USP), em São Paulo (SP), nos dias 06 e 07 de dezembro de 2022, e no Encontro do GT de Fonética e Fonologia realizado no âmbito do 37º Encontro Nacional da ANPOLL (37º ENANPOLL), no período de 03 a 05 de outubro de 2023, na Universidade Federal Fluminense (UFF), em Niterói (RJ). Os organizadores do volume são Ronaldo Manguiera Lima Júnior (UnB - Universidade de Brasília) e Flaviane Romani Fernandes Svartman (USP - Universidade de São Paulo), respectivamente, coordenador e vice-coordenadora do já referido GT no período de realização dos eventos mencionados.

Os trabalhos do presente volume versam sobre temáticas diversas da área de fonética e fonologia do português e de línguas em contato e de línguas estrangeiras adquiridas como língua adicional por falantes nativos do português brasileiro, tais como: aspectos segmentais e suprasegmentais de variedades brasileiras e de línguas africanas em contato com o português, processos fonológicos do português brasileiro, relação entre fonologia e ortografia em português brasileiro, questões de produção e percepção na aquisição de língua estrangeira por falantes nativos de português brasileiro e aspectos fonético-fonológicos, de natureza segmental e suprasegmental, da fala atípica em português brasileiro. Tais trabalhos são uma amostra representativa do estado da arte dos estudos em fonética e fonologia realizados no Brasil e estão dispostos, neste volume, na ordem especificada a seguir.

Abre o volume o texto intitulado *Reestruturação de encontros obstruinte-líquida*, que tem como autora Maria Cantoni. O artigo visa ao estudo das alternâncias entre a estrutura silábica CCV e suas frequentes reestruturações CV.CV e CV no português brasileiro, com base em análise acústica de gravações de fala espontânea de 22 falantes nativos. A pesquisa revela que fatores como a posição do acento influenciam essas reestruturações, evidenciando a maior frequência de CV em contextos não acentuados e de CV.CV em contextos acentuados. Além disso, a qualidade da vogal de apoio, frequentemente centralizada e influenciada pela vogal seguinte, e a duração vocal se mostraram elementos cruciais para a compreensão da dinamicidade e gradualidade do fenômeno. Os resultados contribuem para os debates sobre princípios fonológicos que governam a formação de ataques silábicos.

Em *O alçamento não assimilatório das vogais /e, o/ pretônicas no português do sul do Brasil: uma abordagem fonológica de contraste e aperfeiçoamento*, Elisa Battisti trata do alçamento das vogais médias pretônicas /e, o/ não desencadeado por vogal alta seguinte no português brasileiro de Porto Alegre. O artigo tem como objetivos: (i) buscar explicações para as proporções de alçamento observadas, relativamente menores para /e/ do que para /o/; e (ii) representar os contrastes vocálicos do

português brasileiro em um modelo com traços privativos pouco explorado na literatura de fonologia do português, cuja vantagem consiste em sua estrutura implicacional de traços privativos (gestos), dimensões (contrastes) e níveis superordenados, que explica e restringe a atividade fonológica possível na língua. A análise desenvolvida pela autora revela que a vogal /o/ é mais suscetível ao alçamento porque é menos marcada do que /e/ no sistema de contrastes da língua. A limitação da análise apresentada no artigo, segundo a própria autora, diz respeito à marcação da dimensão *altura da língua* em duas camadas da hierarquia contrastiva para dar conta dos quatro contrastes de altura do português.

O terceiro artigo, intitulado *Os efeitos do aprendizado das letras <e, o> na produção oral de vogais médias pretônicas e postônicas finais*, de autoria de Cecília Toledo, aborda a influência do aprendizado das letras <e, o> na produção de vogais médias e altas pretônicas e postônicas finais na fala da variedade belo-horizontina do português brasileiro. A literatura prévia sobre essa temática documenta que, ao longo da escolarização, a escrita pode retroalimentar a fala. O estudo de Cecília avança em relação à literatura precedente, ao observar os efeitos da escrita na fala em dois contextos acentuais distintos em português brasileiro: pretônico e postônico final. Os resultados quantitativos obtidos pela autora mostraram que os aprendizes aumentam a pronúncia de vogais médias, em relação à pronúncia de vogais altas, em contexto pretônico, mas não em contexto postônico final. Com base em uma análise dos dados à luz dos Modelos de Exemplares (Johnson, 1997; Bybee, 2001; Pierrehumbert, 2001), a autora argumenta que o conhecimento ortográfico pode auxiliar, probabilisticamente, na organização do conhecimento linguístico, contribuindo para o fortalecimento de categorias fonológicas. Entretanto, a ortografia opera de maneira específica em cada categoria, tendo em vista que as categorias fonológicas são múltiplas, probabilísticas e dinâmicas. Respalhada por sua análise de dados, Cecília afirma que a fala e a escrita podem estabelecer múltiplas relações e, não se limitando à ideia unirrepresentacional de que letras representam fonemas, propõe que uma mesma letra – <e, o> – pode estar relacionada a categorias linguísticas múltiplas e probabilísticas.

Por sua vez, *O processamento lexical do acento em pseudopalavras a partir da leitura de falantes do português brasileiro*, de Aline Benevides, examina o papel da similaridade fonológica e da frequência lexical na atribuição de acento em pseudopalavras criadas a partir de palavras reais do português brasileiro. Por meio de um experimento com 34 falantes, foram identificados efeitos da semelhança e da frequência na produção do padrão acentual. Os resultados sugerem que o acento é armazenado lexicalmente e que há interação entre os níveis segmental, silábico e acentual. A pesquisa avança na compreensão dos processos analógicos envolvidos na atribuição do acento em pseudopalavras, destacando a relevância dos padrões segmentais e prosódicos do idioma.

Em *Sândi vocálico externo em português brasileiro: elisão, ditongação e coalescência*, Magnun Madruga e Guilherme Gonçalves apresentam uma análise fonético-acústica do fenômeno de elisão da vogal baixa /a/ no português brasileiro em contextos de sândi vocálico externo. A partir de um estudo experimental com oito falantes, identificaram-se três padrões principais: elisão, ditongo decrescente e coalescência. Este último, inédito na literatura, resulta na emergência de vogais médias [e] ou [o] em sequências específicas. Além de discutir os resultados experimentais, o estudo propõe uma análise fonológica que incorpora a coalescência como subprocesso de elisão e

elimina a necessidade de degeminação, oferecendo uma visão mais econômica e realista das operações fonológicas do PB.

Shirley Freitas e Amanda Balduino descrevem, em *Vogais tônicas orais do kabuverdianu do Príncipe: uma análise fonética*, as características acústicas das vogais orais tônicas do kabuverdianu falado na ilha do Príncipe a partir da análise acústica de dados de fala de quatro falantes nativos. Identificaram-se nove fones vocálicos, ampliando as descrições anteriores da língua e revelando um padrão de dispersão acústica relevante para a definição das vogais. As vogais médias apresentam proximidade acústica entre seus pares altos e baixos, destacando-se como um aspecto particular da organização vocálica do kabuverdianu no Príncipe. A pesquisa contribui para o aprofundamento do conhecimento sobre a fonética e a fonologia dessa língua em sua variedade específica.

No segundo artigo sobre línguas africanas em contato com o português, intitulado *O onset no angolano moderno* e constante do volume, Manuele Bandeira e Ubiratã Alves exploram a estrutura silábica do angolano, língua falada em São Tomé e Príncipe. Este estudo analisa 1.524 itens coletados com falantes bilíngues. Identificaram-se diversas estruturas de ataque silábico, com destaque para onsets simples e complexos, que seguem o Princípio da Sequência de Sonoridade. Além disso, nasais em início de palavra com obstruintes foram analisadas como oclusivas pré-nasalizadas, constituídas por um único segmento de contorno em *onset*. A pesquisa contribui para o entendimento das estruturas fonológicas do angolano, reforçando a relevância da língua para estudos tipológicos.

Em *Correlações entre percepção e produção de palavras heterotônicas do espanhol por aprendizes brasileiros*, Pollianna Millan investiga se a percepção de palavras heterotônicas do espanhol, por aprendizes brasileiros, transforma a produção das mesmas. O estudo realizado pela autora contou com grupos diferentes de participantes: aprendizes brasileiros com menos exposição acadêmica que receberam e que não receberam treinamento perceptual e aprendizes brasileiros com mais exposição acadêmica que receberam e que não receberam treinamento perceptual. Os aprendizes desses diferentes grupos ouviram heterotônicos falados por locutores hispano-falantes e, depois, produziram os mesmos heterotônicos em frases com contexto e em frases-veículo. Os resultados alcançados pela autora, a partir da aplicação de *testes de correlação de Spearman*, demonstraram que há uma tendência maior de correlação entre percepção e produção no grupo com menos exposição acadêmica e que participou do treinamento perceptual.

Já no artigo *Desenvolvimento dos sons vocálicos do inglês como L2 por brasileiros em contexto de imersão*, Anderson Romário Souza Silva, Ronaldo Manguiera Lima Júnior, Lucas Heitor Ananias Oliveira e Dominyk Santos Dias Simões, com base na teoria de língua enquanto sistema dinâmico e complexo, apresentam uma discussão sobre atrito linguístico no português brasileiro (PB) e reorganização do sistema vocálico no inglês como segunda língua (IL2). Os autores analisam o efeito do contexto de imersão na segunda língua para o desenvolvimento do sistema vocálico do IL2 e a reorganização do sistema vocálico do PB, através de uma pesquisa longitudinal com dados referentes a coletas anteriores, no decorrer e posteriores a um período de estadia de nove meses de estudantes brasileiros nos Estados Unidos da América. São analisados valores de F1 e F2 e de duração de vogais nos dados das duas línguas produzidos pelos falantes que participaram das coletas. Conforme os resultados alcançados, houve reorganização

das vogais do inglês-L2, mas não houve mudanças consideráveis nas vogais do PB.

Geovana Soncin, Cecília Lorena Silva Guida, Fernanda Leitão de Castro Nunes de Lima e Larissa Cristina Berti, no artigo *Diferenças na produção de foco prosódico contrastivo na fala de adultos e de crianças com aquisição fonológica típica e atípica do Português Brasileiro*, têm como objetivo o estudo comparativo da caracterização acústica da marcação de foco prosódico em dados de fala de português brasileiro produzidos por adultos, crianças com desenvolvimento típico de linguagem e crianças com transtorno fonológico nativos dessa variedade de português. Os dados abordados no estudo foram obtidos por meio de experimentos elaborados para a produção de sentenças com foco prosódico contrastivo e sentenças neutras. Nesses dados produzidos pelos três grupos distintos de falantes, foram analisados os seguintes parâmetros acústicos: frequência fundamental, duração e intensidade. Os resultados alcançados indicaram performance distinta dos grupos na marcação de foco contrastivo na análise dos três parâmetros fonéticos e sugerem que crianças com transtorno fonológico podem apresentar instabilidades no plano prosódico.

O artigo que fecha o volume, intitulado *Os impactos vocais da doença de Parkinson em idosos*, de autoria de Lucas Manca Dal’Ava e Plínio Almeida Barbosa, visou à análise das diferenças nos parâmetros acústicos de dimensões fonéticas diversas em discursos espontâneos e de leitura em voz alta realizados por idosos brasileiros com envelhecimento saudável e com doença de Parkinson (DP). O objetivo do estudo foi avaliar o impacto da doença de Parkinson (DP) e do envelhecimento saudável nos biomarcadores de fala de idosos brasileiros, incluindo parâmetros acústico-prosódicos. Foram analisadas gravações de fala de idosos brasileiros nativos, considerando três parâmetros acústico-prosódicos: frequência fundamental, duração e intensidade. Segundo os autores, os resultados obtidos da análise acústica dos dados indicam que certos parâmetros prosódico-acústicos, como F0, ênfase espectral, taxas de articulação e elocução e *shimmer*, podem ser afetados devido às restrições motoras da DP. Os autores ainda destacam a necessidade de estudos posteriores longitudinais e com amostras mais diversas para uma melhor compreensão dos efeitos da DP e do envelhecimento vocal.

Este volume reflete a riqueza e a pluralidade dos estudos em Fonética e Fonologia, abarcando temas que vão desde processos fonológicos específicos, segmentais e prosódicos, até interfaces com a psicolinguística, a aquisição de segunda língua e a saúde vocal. As contribuições dialogam com questões fundamentais da linguística teórica e aplicada, ao mesmo tempo que demonstram o impacto prático dessas investigações em áreas como a educação, a terapia da fala e a análise de línguas minoritárias. Ao reunir pesquisas que exploram diferentes fenômenos e perspectivas, este volume reafirma a relevância da Fonética e da Fonologia para o desenvolvimento científico da Linguística no Brasil, consolidando o papel dessas áreas na compreensão da linguagem humana em suas múltiplas dimensões.

Ronaldo Manguiera Lima Júnior, *Universidade de Brasília*

Flaviane Romani Fernandes Svartman, *Universidade de São Paulo*

## Reestruturação de encontros obstruinte-líquida<sup>1</sup>

### *Restructuring of obstruent-liquid clusters*

Maria Mendes Cantoni\*

*Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil*

**Resumo:** No português, sílabas CCV formam encontros de obstruinte-líquida e podem ser opcionalmente reestruturadas pela (1) inserção de uma vogal entre as consoantes do ataque, resultando em CV.CV, ou (2) o apagamento da consoante líquida na segunda posição, resultando em CV, casos que foram avaliados separadamente em estudos anteriores. Este estudo apresenta uma avaliação integrada da alternância CCV~CV.CV~CV no português brasileiro, a partir de análise acústica das ocorrências de CCV em gravações de fala espontânea de 22 falantes nativos. Buscou-se determinar o papel de fatores linguísticos, como a posição do acento, bem como avaliar a qualidade da vogal de apoio e padrões de duração decorrentes da perda consonantal ou da inserção vocálica. Em termos de sua qualidade, a vogal de apoio é uma vogal centralizada, podendo haver harmonia com a vogal seguinte. Verificou-se que o padrão CV é mais frequente em contexto não acentuado e que leva ao alongamento da vogal, enquanto o padrão CV.CV é mais frequente em contexto acentuado. Os resultados são discutidos em termos da gradualidade e dinamicidade do fenômeno e considerando princípios fonológicos de reestruturação silábica e formação de ataques.

**Palavras-chave:** Encontros consonantais. Vogal epentética. Português brasileiro. Fala espontânea. Modelos de Uso.

**Abstract:** In Portuguese, CCV syllables form obstruent-liquid clusters and can be optionally restructured by (1) inserting a vowel between the onset consonants, resulting in CV.CV, or (2) deleting the liquid consonant in the second position, resulting in CV, cases that were evaluated separately in previous studies. This study presents an integrated evaluation of the CCV~CV.CV~CV alternation in Brazilian Portuguese, based on acoustic analysis of CCV occurrences in spontaneous speech recordings from 22 native speakers. We sought to determine the role of linguistic factors, such as the position of the accent, as well as evaluate the quality of the epenthetic vowel and duration patterns resulting from consonant loss or vowel insertion. In terms of its quality, the supporting vowel is a centralized vowel, and there may be harmony with the following vowel. It was found that the CV pattern is more frequent in a non-accented context and leads to vowel lengthening, while the CV.CV pattern is more frequent in an accented context. The results are discussed in terms of the gradualness and dynamicity of the phenomenon and considering phonological principles of syllabic restructuring and onset formation.

**Keywords:** Consonantal cluster. Epenthetic vowel. Brazilian Portuguese. Spontaneous Speech. Usage-based models.

---

<sup>1</sup> Agradecimentos: A autora agradece aos membros da equipe do Laboratório de Fonologia da UFMG que contribuíram com a segmentação e anotação dos dados: Bruna Egito, Henrique Nevado, Júlia Viegas, Paula Silva. Agradece também pelos comentários recebidos durante a apresentação de uma versão prévia do trabalho no encontro do GT de Fonética e Fonologia no 37o Enanpoll, e também pelos comentários de Magnun Rochel Madruga e de dois avaliadores anônimos. Este trabalho recebeu apoio financeiro da Fapemig, por meio do edital Universal, projeto APQ 02997-18.

\* Professora adjunta, Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil; [mmcantoni@gmail.com](mailto:mmcantoni@gmail.com)

## 1 INTRODUÇÃO

No português, sílabas CCV ocorrem na forma de encontros de obstruente-líquida (EOL), em que a obstruente na primeira posição do ataque pode ser [p, b, t, d, k, g, f, v] e a líquida na segunda posição pode ser [r, l]. As sílabas CCV ocorrem em uma parcela pequena das palavras da língua (4,4% contra 60,6% para as sílabas CV, segundo Viaro e Guimarães-Filho 2007) e há lacunas em sua distribuição (e.g. algumas combinações não ocorrem em início de palavra, como [vr] e [tl]). Além disso, alguns encontros são raros ou restritos a palavras estrangeiras (e.g. [vl]) e os encontros com tepe são mais frequentes do que os com lateral. Trata-se, portanto, de um tipo silábico de baixa frequência e com restrições de ocorrência que, como abordaremos neste estudo, pode alternar com o tipo mais frequente e abrangente CV.

No português brasileiro, os EOLs podem ser opcionalmente reestruturados (1) pela inserção de uma vogal entre as consoantes do ataque, ou (2) pelo apagamento de elementos do encontro. O primeiro caso, ilustrado no Exemplo 1, resulta em duas sílabas, CV.CV, e pode ser entendido como um processo fonológico de fortalecimento (cf. Stampe, 1979; Donegan; Stampe, 2009), por envolver a inserção de um som. A vogal epentética é tradicionalmente referida como vogal de apoio, para diferenciá-la de outro caso de epêntese observado na língua e que envolve encontros heterossilábicos (e.g. pacto ['paktʊ]~['paktʊ]). Em estudos anteriores, foi observado que a vogal inserida é neutra ou em harmonia com a vogal da sílaba CCV (Silva, 1996; Nishida, 2005; Silveira; Seara, 2017).

(1) Epêntese vocálica (vogal de apoio): fortalecimento

CVCV [ˈpratʊ]~[pəˈratʊ]~[paˈratʊ] *prato*

Já o segundo caso, ilustrado no Exemplo 2, por outro lado, pode ser considerado um caso de lenição (cf. Stampe, 1979; Donegan; Stampe, 2009), pois envolve o apagamento de sons. Geralmente ocorre a omissão da consoante líquida na segunda posição, como em 2a, resultando em CV. Pode também ser apagada a obstruente, como 2b, resultando no mesmo tipo silábico. Nesses casos, estudos anteriores mostraram que pode haver alongamento compensatório da vogal com a simplificação do ataque (Cristófar-Silva 2000, 2003; Cristófar-Silva et al., 2017; Cristófar-Silva; Almeida, 2008; Oliveira, 2017). É possível ainda que a vogal seja apagada, mas não as consoantes, como em 2c, ou que tanto a vogal quanto a líquida sejam apagadas, como em 2d. Nesses dois últimos casos, não é trivial determinar qual o tipo silábico formado, que pode envolver uma consoante silábica ou o preenchimento da coda da sílaba precedente.

(2) Apagamento: lenição (Cristófar-Silva et al., 2017)

- |    |      |                              |                   |
|----|------|------------------------------|-------------------|
| a. | C1V  | [pɾi.si.ˈpaw]~[pɾi.si.ˈpaw]  | <i>principal</i>  |
| b. | C2V  | [kõs.tru.ˈsãʊ]~[kõs.ru.ˈsãʊ] | <i>construção</i> |
| c. | C1C2 | [ˈo.trʊ]~[ˈo.tr]             | <i>outro</i>      |
| d. | C1   | [ˈlitro]~[ˈlit]              | <i>litro</i>      |

Os padrões descritos nos Exemplos 1 e 2 (a, b), em conjunto, geram na língua uma alternância de tipos silábicos: CCV ~ CV ~ CVCV. Em estudos anteriores, os dois casos foram avaliados separadamente, impedindo o entendimento geral das

motivações das alternâncias. Este estudo apresenta uma avaliação integrada da reestruturação de encontros obstruinte-líquida no português brasileiro, ao considerar que os padrões sonoros da língua estão interligados por associações probabilísticas (Pierrehumbert, 2001, 2003, 2022) e que o conhecimento gramatical é construído pelo uso linguístico (Bybee, 2001, 2006). As seguintes perguntas serão abordadas: Qual a incidência de apagamento e de epêntese em EOL na fala? Quais fatores favorecem o apagamento e a epêntese? Qual a qualidade fonética da vogal de apoio? Como os dois fenômenos se relacionam?

Como iremos mostrar, a alternância  $CCV \sim CV \sim CVCV$  é dinâmica e foneticamente motivada:  $CVCV$  predomina em sílabas tônicas, cujo alongamento acentual propicia duração suficiente para a emergência de uma vogal adicional, enquanto  $CV$  predomina nas átonas, cujo encurtamento favorece a supressão parcial ou total de elementos sonoros e segue padrões probabilísticos da língua. Ademais, a alternância é foneticamente gradual, pois envolve uma progressão de duração entre os três padrões silábicos envolvidos.

## 2 METODOLOGIA

O presente estudo procurou avaliar a reestruturação de EOLs usando gravações de fala espontânea, obtidas da base de dados CEFALA-1 (Follador Neto; Silva; Yehia, 2019), que contém gravações de um total de 104 falantes coletadas em cabine acústica e equipamento profissional. Foi selecionada aleatoriamente uma amostra de 22 falantes (12 F, 10 M) e o áudio obtido pelo microfone M1, de lapela. Foram utilizados os trechos de fala espontânea, com uma média de 2 minutos de duração. Todos os falantes selecionados são originários de Minas Gerais, sem indícios de desvios de fala e com idade aparente entre 20 e 40 anos, segundo avaliação independente de três pesquisadores que colaboraram com o presente estudo.

Os procedimentos de avaliação dos dados seguiram etapas de processamento, medição e análise quantitativa. As etapas de processamento e de medição foram realizadas no *software* Praat (Boersma; Weenink, 2023), e contou com a atuação de uma equipe de segmentadores treinados e revisão pela autora. A análise quantitativa foi realizada na linguagem R (R Core Team, 2023).

O processamento dos dados teve início com a transcrição ortográfica manual dos trechos de fala espontânea. Em seguida, foi feita uma busca exaustiva pelos encontros de obstruinte e líquida ortográficos, que expressam a forma canônica da alternância  $CCV \sim CV \sim CVCV$  e totalizaram 321 ocorrências. O apêndice A mostra algumas imagens ilustrativas dos padrões encontrados e sua classificação. Foi feita, então, a segmentação e a anotação manual dos encontros e da vogal de apoio, quando ocorreu. A segmentação dos sons seguiu critérios usuais em fonética acústica (Turk et al., 2006) e considerando ainda as características dos sons do português brasileiro (Cristófar-Silva et al., 2019). Vogais de foram identificadas pela presença de formantes e delimitadas pelo início e fim de F2. Tepes foram identificados como um trecho breve (da ordem de 20-30 ms) com baixa amplitude e escassez de energia no espectrograma. Laterais foram identificadas pela presença de formantes mais fracos que os de uma vogal e na região de 1000-1500 Hz e 2500-3000 Hz. Oclusivas foram identificadas como um trecho de baixa amplitude maior do que 40 ms, delimitadas à esquerda pela queda de amplitude do fechamento e à direita pelo transiente da soltura,

se visível. Fricativas foram identificadas como um intervalo contendo sinal aperiódico contínuo.

Na etapa de medição, por meio de *script*, foi levantado o número de ocorrências dos encontros e medidos a duração do encontro e da vogal de apoio e os dois primeiros formantes (F1 e F2, respectivamente associados à altura e à anterioridade vocálica) da vogal de apoio.

Na avaliação quantitativa, foram ajustados modelos lineares mistos para a variável de resposta duração, e generalizados para a variável resposta tipo de padrão silábico, por meio das funções *lmer* e *glmer* do pacote *lme4* (Bates et al., 2015). Nos dois casos, os modelos foram construídos com efeitos aleatórios de intercepto para participante e palavra. Para a modelagem dos padrões silábicos, os efeitos fixos de padrão acentual e sexo foram usados para construir quatro modelos diferentes: contendo ambos os efeitos, com interações ou sem interações, ou contendo apenas um dos efeitos. Os quatro modelos foram comparados par-a-par, por meio da função *anova*, selecionando-se o modelo mais complexo para o qual houvesse uma diferença significativa. O modelo de melhor ajuste contém os efeitos fixos de acento e sexo sem interações e obteve AIC = 236,5. O ajuste do modelo foi seguido de teste II qui-quadrado de Wald por análise de variância com a função *Anova* do pacote *car* (Fox; Weisberg, 2019), de forma a avaliar a contribuição de cada efeito separadamente, e testes de comparações múltiplas de médias com contraste de Tukey, com a função *glht* do pacote *multcomp* (Hothorn et al., 2008), para avaliar a significância dos níveis dos efeitos fixos. Pela inspeção de gráficos *qq*, checkou-se que os resíduos do modelo ajustado não violam a assunção de normalidade. Para a modelagem da duração dos EOL e da duração das vogais de apoio, além das variáveis padrão acentual da sílaba e sexo do falante, foram usadas como efeitos fixos também as variáveis vozeamento da consoante C1 do EOL, vogal seguinte e padrão silábico realizado. Para cada tipo de duração, o melhor modelo foi obtido por procedimento análogo ao descrito no caso anterior, mas usando a função *buildmer* do pacote de mesmo nome (Cesko, 2020) para automatizar o procedimento de comparação de modelos. O modelo máximo não incluiu interações, por motivo de falta de convergência, e os efeitos aleatórios foram mantidos durante todo o procedimento de comparação. O modelo de melhor ajuste para a duração dos EOL contém os efeitos fixos de padrão silábico, vozeamento, acento e vogal e obteve AIC = -1143. O modelo de melhor ajuste para a duração das vogais de apoio contém os efeitos fixos de acento e obteve AIC = -956. Após o ajuste do melhor modelo, os mesmos procedimentos de avaliação usados no caso anterior foram seguidos.

### 3 RESULTADOS

#### 3.1 Tipo de líquida

Do total de 321 encontros CCV canônicos, foram encontrados 32 encontros com lateral e 289 encontros com tepe, como mostra a Tabela 1. Encontros com lateral não apresentaram ocorrência de epêntese (i.e. vogal de apoio) e contaram com menor índice de apagamento do que os encontros com tepe ( $\chi^2 = 3,08$ ;  $df = 1$ ;  $p\text{-value} = 0,0792$ ). Tais resultados podem se dever ao fato de que a lateral apresenta características vocálicas, mas também pode ser uma consequência da menor frequência

de ocorrência das palavras que contêm EOL com lateral, em comparação com as palavras que contêm EOL com tepe.

Tabela 1 - Padrões de alternância de CCV por tipo de consoante líquida.

	Apagamento	CCV	Epêntese	Total
Lateral	5	27	0	32
Tepe	45	92	152	289

Fonte: Elaborado pela autora.

Os demais resultados apresentados a partir deste ponto se concentrarão nos EOL com tepe, apenas. Nesses, prevalece a ocorrência de vogal de apoio, que corresponde a 52,6% dos dados, contra 31,8% para CCV e 15,6% para algum tipo de redução.

### 3.2 Tipo de obstruente

A Tabela 2 apresenta os padrões de alternância observados para cada tipo de obstruente na primeira posição de CCV. A discrepância no número total de ocorrências de algumas consoantes ([d, f, v]) em CCV reflete distribuições da língua e limita a avaliação estatística da alternância. Observa-se que [p, b, t, g] apresentam maiores índices de epêntese do que dos demais padrões.

Tabela 2 - Padrões de alternância de CCV por tipo de obstruente.

	Apagamento	CCV	Epêntese	Total
p	12	39	48	99
b	4	5	18	27
t	22	28	45	95
d	3	1	4	8
k	1	12	8	21
g	2	3	22	27
f	0	3	4	7
v	2	1	2	5

Fonte: Elaborado pela autora.

### 3.3 Vogal seguinte ao EOL

A Tabela 3 mostra os padrões de alternância conforme a vogal que ocorre após o EOL – ou seja, a vogal V da sílaba canônica CCV. As contagens das vogais médias de mesma anterioridade [e, ε] e [o, ɔ] foram fundidas, uma vez que nem sempre é possível prever ortograficamente a altura de tais vogais. Observa-se que as vogais não altas, não posteriores [e, ε, a] apresentam maiores valores de epêntese do que apagamento e CCV, ao passo que nas demais vogais os índices de CCV e epêntese são equivalentes.

Tabela 3 - Padrões de alternância de CCV por vogal após o EOL.

	Apagamento	CCV	Epêntese	Total
i	9	19	19	47
e, ε	11	20	35	66
a	11	16	58	85
o, ɔ	12	36	38	86
u	3	1	1	5

Fonte: Elaborado pela autora.

### 3.4 Tonicidade da sílaba

Considerando os padrões de alternância em relação à tonicidade da sílaba canônica CCV, como mostrado pela Tabela 4, observa-se que a vogal de apoio ocorre mais nas tônicas do que nas pretônicas e nas postônicas. Ademais, observa-se que o apagamento afeta mais as postônicas que as pretônicas. A modelagem estatística dos padrões de alternância é apresentada na seção 3.6 e suporta a diferença entre grupos apontada acima.

Esse padrão de ocorrência é previsível, uma vez que as sílabas tônicas no português se caracterizam pela incidência de fenômenos de fortalecimento, pela preservação segmental e pelo contraste máximo no inventário vocálico, enquanto as átonas, especialmente as postônicas, estão sujeitas a fenômenos de lenição, incluindo apagamentos dos mais diversos tipos, e apresentam simplificação de contrastes no inventário de vogais (Bisol, 2010; Cristófaros-Silva, 2002). Esse cenário está em conformidade com as funções clássicas dos processos fonológicos de fortalecimento e lenição<sup>2</sup> e justifica a motivação fonética da alternância.

Tabela 4 - Padrões de alternância de CCV por tonicidade da sílaba.

	Apagamento	CCV	Epêntese	Total
Tônica	6	18	66	90
Pretônica	17	52	53	122
Postônica	22	22	33	77

Fonte: Elaborado pela autora.

### 3.5 Sexo do falante

Como se pode ver na Tabela 5, a epêntese de vogal de apoio é a manifestação que predomina no grupo das mulheres e ocorre mais neste grupo do que entre os homens. À parte as interpretações sociolinguísticas que possa ter, o que se pretende

<sup>2</sup> “Fortitions apply to enhance the divisions of the prosodic score and the clarity of intended sounds, and then Lenitions apply to enhance the fluency of the sequence of sounds. Fortitions make sounds more like themselves and less like adjacent sounds; lenitions make sounds more like their adjacent sounds, sometimes to the point of identity.” (Donegan, Stape, 2009, p. 2).

com esse resultado é usá-lo para ancorar a próxima subseção, ao se tratar da qualidade da vogal de apoio. A modelagem estatística dos padrões de alternância é apresentada na seção 3.6 e suporta a diferença entre grupos apontada acima.

Tabela 5 - Padrões de alternância de CCV por sexo do falante.

	Apagamento	CCV	Epêntese	Total
Mulheres	12	36	94	142
Homens	33	56	58	147

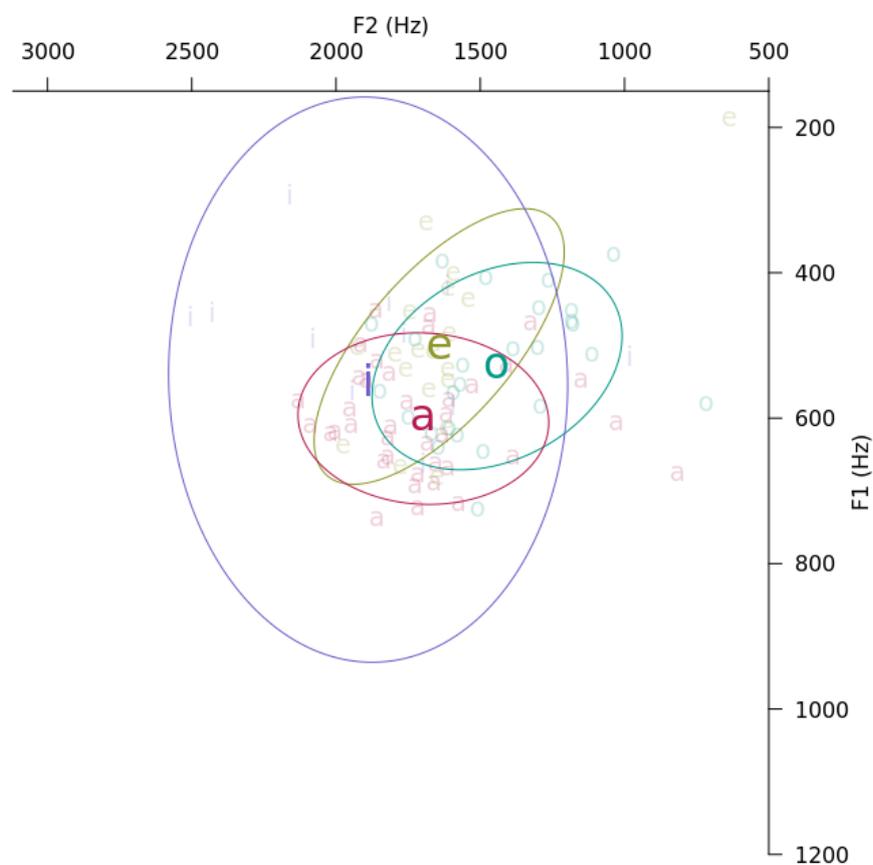
Fonte: Elaborado pela autora.

### 3.6 Modelagem da alternância de padrões silábicos

O modelo linear misto generalizado melhor ajustado aos dados da variável de padrões silábicos tem como efeitos fixos as variáveis preditoras padrão acentual e sexo do falante. O teste II qui-quadrado de Wald confirmou que tanto padrão acentual ( $\chi^2 = 11,3681$ ;  $df = 2$ ;  $p = 0,003400$ ) quanto sexo ( $\chi^2 = 8,0627$ ;  $df = 1$ ;  $p = 0,004519$ ) são relevantes para o modelo. Os testes de comparações múltiplas indicaram que, entre os níveis da variável padrão acentual, é significativa a diferença entre postônica e tônica, estimada em -1,7522 com erro padrão de 0,5412 ( $z = -3,238$ ;  $p = 0,00361$ ) e também a diferença entre postônica e pretônica, estimada em -1,0819, com erro padrão de 0,4539 ( $z = -2,384$ ;  $p = 0,03428$ ). Entre os níveis da variável sexo, também foi significativa a diferença entre os níveis homem e mulher, estimada em -1,5274, com erro padrão de 0,5379 ( $z = -2,839$ ;  $p = 0,00452$ ).

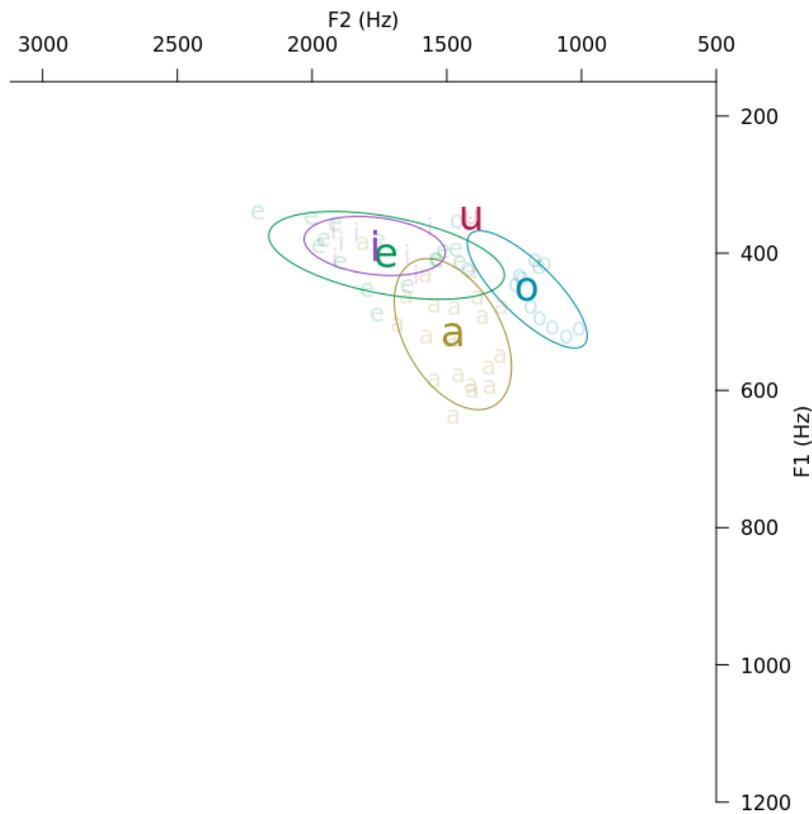
### 3.7 Qualidade da vogal de apoio

O espaço acústico, composto pelos valores de F1 e F2 das vogais de apoio para mulheres e homens, é apresentado nos Gráficos 1 e 2, respectivamente. As ocorrências foram identificadas e agrupadas com o símbolo da vogal que segue o EOL. As dimensões do gráfico foram estabelecidas de forma a comportar valores típicos para qualquer vogal do português, permitindo avaliar o posicionamento das vogais de apoio nesse espaço. Como se pode observar, a vogal de apoio (F1 com mediana de 505 Hz e F2 com mediana de 1615 Hz) é frequentemente uma vogal centralizada, considerando os valores típicos dos formantes para as vogais do português.



Fonte: Elaborado pela autora.

Gráfico 1 - Espaço acústico das vogais de apoio das mulheres.



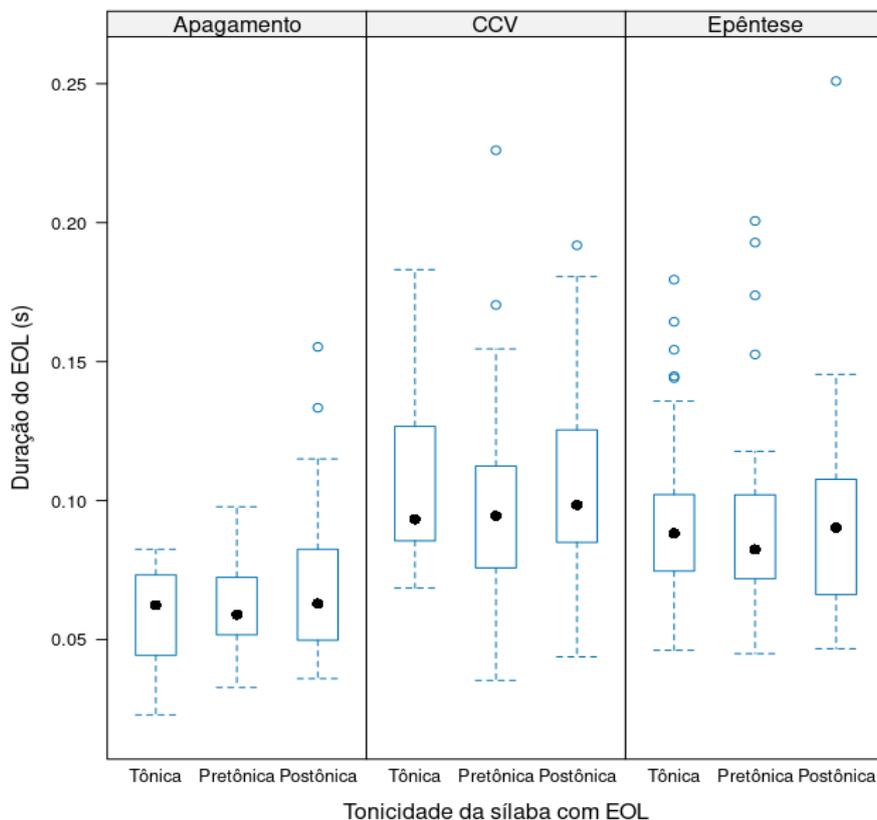
Fonte: Elaborado pela autora.

Gráfico 2 - Espaço acústico das vogais de apoio dos homens.

Os dados foram separados de acordo com o sexo por duas razões: pela diferença dos valores de formantes típicos desses grupos, decorrentes de fatores anatômicos, e pela diferença nos padrões linguísticos observados entre os grupos. No grupo dos homens, a vogal de apoio tende a seguir a qualidade da vogal do EOL, mas não no grupo das mulheres. Retomando as taxas de ocorrência de cada padrão de alternância por grupo, mostrados na Tabela 5, observou-se que as mulheres apresentam maiores índices de vogais de apoio que os homens. Dessa forma, a alta repetição do padrão de epêntese – pelas mulheres – estaria relacionada com uma desassociação entre a qualidade da vogal de apoio e a vogal do EOL. Considerando que a epêntese é foneticamente motivada, pois favorecida em sílabas tônicas, é natural esperar que a vogal de apoio seja uma antecipação da vogal seguinte ao EOL, pelo alongamento dos gestos usados para produção da sílaba CCV. Nesse cenário, a centralização da vogal de apoio pode ser entendida como um processo adicional de lenição, gerenciada pela frequência de uso.

### 3.8 Padrões de duração dos EOL e da vogal de apoio

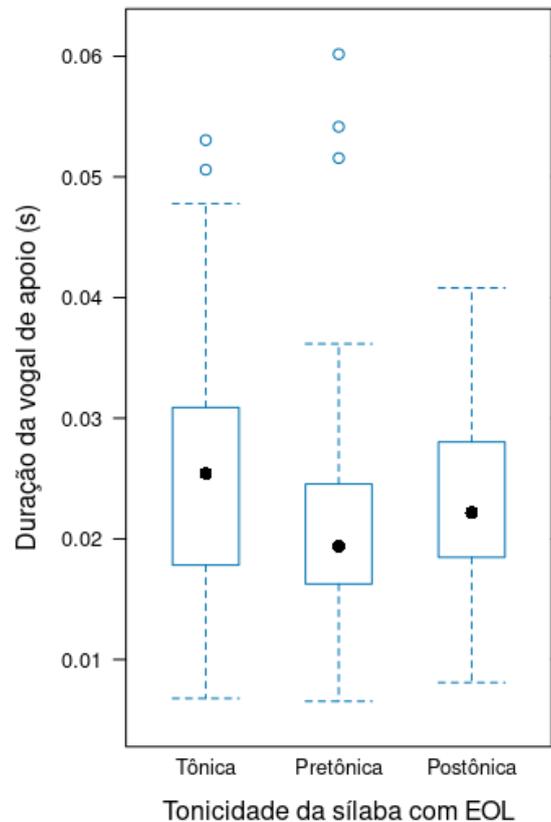
A duração dos EOL, contemplando apenas as consoantes e descontando a duração da vogal de apoio, quando ocorreu, é mostrada no Gráfico 3. A modelagem estatística dos padrões é apresentada na seção 3.9 e suporta as avaliações realizadas a seguir. O padrão envolvendo apagamento apresentou valores de duração menores do que CCV e do que epêntese, e CCV apresentou valores menores do que epêntese.



Fonte: Elaborado pela autora.

Gráfico 3 - Duração dos EOL por padrão de alternância e por nível de tonicidade da sílaba.

A duração das vogais de apoio agrupadas pela tonicidade da sílaba do EOL é mostrada no Gráfico 4. Vogais que ocorrem nas sílabas tônicas são mais longas do que as que ocorrem em pretônicas, contudo não houve diferença entre tônicas e postônicas e entre postônicas e pretônicas.



Fonte: Elaborado pela autora.

Gráfico 4 - Duração das vogais de apoio por nível de tonicidade da sílaba.

Em conjunto, os valores de duração indicam que os EOL e as vogais de apoio são afetados pela tonicidade da sílaba, havendo maior duração nas tônicas do que nos demais padrões, dando suporte à interpretação de motivação fonética pelo alongamento acentual apontada anteriormente na seção 3.4.

### 3.9 Modelagem dos padrões de duração

O modelo linear misto melhor ajustado aos dados da variável duração dos EOL tem como efeitos fixos as variáveis preditoras padrão silábico, nível de vozeamento da consoante C1 do EOL, padrão acentual e vogal seguinte. O teste II qui-quadrado de Wald confirmou que todas as variáveis são relevantes para o modelo:

- padrão silábico ( $\chi^2 = 39,6262$ ;  $df = 2$ ;  $p = 2,485 \times 10^{-9}$ )
- vozeamento ( $\chi^2 = 25,2119$ ;  $df = 1$ ;  $p = 5,136 \times 10^{-7}$ )
- acento ( $\chi^2 = 7,5160$ ;  $df = 2$ ;  $p = 0,02333$ )
- vogal ( $\chi^2 = 9,8749$ ;  $df = 4$ ;  $p = 0,04259$ )

Os testes de comparações múltiplas indicaram que, entre os níveis da variável padrão silábico, é significativa a diferença entre CCV e apagamento, estimada em 0,0298758 com erro padrão de 0,0051879 ( $z = 5,759$ ;  $p = 2,11 \times 10^{-8}$ ) e também a diferença entre epêntese e apagamento, estimada em 0,0291290, com erro padrão de

0,0050310 ( $z = 5,790$ ;  $p = 2,11 \times 10^{-8}$ ). Entre os níveis da variável vozeamento, foi significativa a diferença entre os níveis desvozeada e vozeada, estimada em  $-0,023182$ , com erro padrão de  $0,004617$  ( $z = 5,021$ ;  $p = 5,14 \times 10^{-7}$ ). Entre os níveis da variável acento, foi significativa a diferença entre os níveis postônica e pretônica, estimada em  $-0,013386$ , com erro padrão de  $0,004886$  ( $z = 2,740$ ;  $p = 0,0184$ ). Quanto à variável vogal, apenas duas diferenças foram significativas (i vs. a e i vs. e) e, como de pouco interesse para o estudo, não serão exploradas.

No caso da duração da vogal de apoio, o melhor modelo ajustado é um modelo linear misto que tem como efeito fixo apenas a variável preditora padrão acentual. O teste II qui-quadrado de Wald confirmou a variável é relevante para o modelo ( $\chi^2 = 7,1288$ ;  $df = 2$ ;  $p = 0,02831$ ). O teste de comparações múltiplas indicou que, entre os níveis da variável padrão acentual, é significativa a diferença entre pretônica e tônica, estimada em  $-0,004553$  com erro padrão de  $0,001723$  ( $z = -2,642$ ;  $p = 0,0247$ ).

#### 4 CONSIDERAÇÕES FINAIS

O presente estudo procurou avaliar a reestruturação de EOLs de forma abrangente, em uma visão integrada da alternância  $CCV \sim CV \sim CVCV$ , ancorada em uma visão da fonologia como probabilística e baseada no uso. Tal perspectiva teórica propiciou suporte à avaliação de como padrões acústico-articulatórios da fala podem desencadear padrões de alternância. Mostrou-se que a alternância nos EOL é foneticamente motivada e que as diferentes combinações de sons nas sílabas – i.e. os padrões fonotáticos da língua – refletem mecanismos da produção (Albano, 2007), resultados que são compatíveis com os pressupostos de que os padrões sonoros da língua estão interligados por associações probabilísticas (Pierrehumbert, 2001, 2003, 2022) e que o conhecimento gramatical é construído pelo uso linguístico (Bybee, 2001, 2006).

As perguntas levantadas inicialmente foram respondidas da seguinte forma, em relação aos encontros  $CCV$  com tepe.

- a) *Qual a incidência de apagamento e de epêntese em EOL na fala?* Foi encontrada uma taxa geral de 52,6% de epêntese ( $CVCV$ ), contra 31,8% para  $CCV$  e 15,6% para apagamento ( $CV$ ).
- b) *Quais fatores favorecem o apagamento e a epêntese?* O apagamento é favorecido em sílabas átonas, enquanto a epêntese é favorecida em sílabas tônicas, em consonância com o esperado para lenição e fortalecimento, respectivamente. A relação com a acentuação (e alongamento/encurtamento) na língua aponta para a motivação fonética dos dois padrões não canônicos da alternância.
- c) *Qual a qualidade fonética da vogal de apoio?* A vogal de apoio pode ser uma vogal semelhante à vogal que segue o EOL (apenas fortalecimento) ou centralizada (fortalecimento pela epêntese e lenição pela centralização).
- d) *Como os dois fenômenos se relacionam?* O apagamento e a epêntese se mostram complementares em termos de sua natureza fonológica (lenição vs. fortalecimento), sendo que no segundo caso pode haver ainda a incidência de outro tipo de lenição, com a centralização da vogal. Ambos são foneticamente motivados e envolvem gradiência em sua implementação, refletindo o impacto do uso linguístico.

Tais questões estão em conformidade com os pressupostos teóricos de que restrições fonotáticas gradientes, que envolvem detalhes fonéticos finos, fazem parte do conhecimento fonológico dos falantes da língua, que é construído a partir da frequência de uso e da probabilidade de ocorrência de padrões sonoros (Pierrehumbert, 2001, 2003, 2022; Bybee, 2001).

Em relação ao apagamento, não foi feita distinção entre as quatro possibilidades apresentadas inicialmente (Exemplo 2a-2d). Contudo, todas foram atestadas e indicam a instabilidade presente nesse padrão de lenição, decorrente da gradiência do fenômeno.

Há também gradiência no percurso de fortalecimento, como indicado pelas associações entre padrões de duração dos elementos da sílaba. Pode ser levantada uma possibilidade de explicar o surgimento da vogal epentética, que ocorre frequentemente em sílabas tônicas com oclusivas, e envolveria uma sequência de ações incrementais e progressivas: a incidência em sílaba tônica leva ao aumento da duração dos sons, com aumento da soltura das oclusivas, que pode então se tornar vozeada e servir de âncora para o surgimento da vogal de apoio. Estudos futuros podem avaliar essa possibilidade de explicação, pela investigação detalhada e experimental de encontros CCV e observação da duração de intervalos subfonêmicos.

## REFERÊNCIAS

- Albano EC. Codificação estatística das categorias fonéticas: vestígio da dinâmica da fala na fonotaxe lexical. *Revista de Estudos da Linguagem*. 2007;15(2):9-42. [citado 06 jun. 2023]. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/2447/2401>.
- Bates D, et al. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. 2015;67(1):1-48. Doi:10.18637/jss.v067.i01.
- Bisol L. A simetria no sistema vocálico do português brasileiro. *Linguística - Revista de Estudos Linguísticos da Universidade do Porto*. 2010;5(1):41-52.
- Boersma P, Weenink D. Praat: doing phonetics by computer [Programa de computador]; 2023. Versão 6.2.03. [citado 05 mai. 2023]. Disponível em: <http://www.praat.org/>.
- Bybee J. *Phonology and language use*. Cambridge: Cambridge University Press; 2001.
- Bybee J. From usage to grammar: the mind's response to repetition. *Language*. 2006;82(1):711-733.
- Cesko CV. Using 'Buildmer' to automatically Find & Compare Maximal (Mixed) models. 2020. Disponível em: <https://cran.r-project.org/web/packages/buildmer/vignettes/buildmer.html>.
- Cristófaros-Silva T. Sobre a queda de encontros consonantais no português brasileiro. *Estudos Linguísticos*. 2000;29(1):522-527.
- Cristófaros-Silva T. *Fonética e fonologia do português: roteiro de estudos e guia de exercícios*. 6. ed. São Paulo: Contexto; 2002.

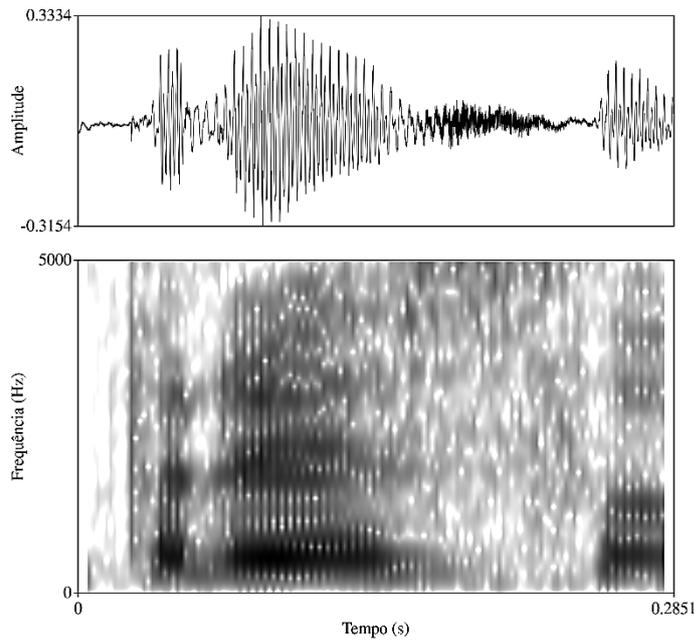
- Cristófaros-Silva T. Sound change in tautosyllabic consonantal clusters in Brazilian Portuguese. Proceedings of the 15<sup>th</sup> International Congress of Phonetic Sciences, Barcelona - Espanha. 2003; p. 1675-1678.
- Cristófaros-Silva T, Almeida LS. On the nature of epenthetic vowels. In: Bisol L, Bresscancini C, organizadoras. Contemporary phonology in Brazil. Cambridge Scholars Publishing; 2008. p. 193-212.
- Cristófaros-Silva T, et al. Fonética acústica: os sons do português brasileiro. São Paulo: Contexto; 2019.
- Cristófaros-Silva T, et al. Segmental loss and phonological representation. In: Lindsey G, Nevins A, organizadores. Sonic signatures. Amsterdam/Philadelphia: John Benjamins Publishing Company; 2017. p. 215-230.
- Donegan P, Stampe D. Hypotheses of natural Phonology. Poznań Studies in Contemporary Linguistics. 2009;45(1):1-31.
- Follador Neto A, Silva AP, Yehia HC. Corpus CEFALA-1: Base de dados audiovisual de locutores para estudos de biometria, fonética e fonologia. Revista de Estudos da Linguagem. 2019;27(1):191-212. [citado 06 jun. 2023]. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/relin/article/view/13378>.
- Fox J, Weisberg S. An R companion to applied regression, 3. ed. Sage: Thousand Oaks; 2019.
- Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biometrical Journal. 2008;50(3):346-363.
- Nishida G. Análise acústica do Tap em grupos no PB. [monografia de conclusão de curso]. Curitiba: Universidade Federal do Paraná; 2005.
- Oliveira NA. Variação em encontros consonantais tautosilábicos no português brasileiro [dissertação]. Belo Horizonte: Universidade Federal de Minas Gerais; 2017.
- Pierrehumbert J. Exemplar dynamics: word frequency, lenition and contrast. In: Bybee JL, Hopper P, editores. Frequency and the emergence of linguistic structure. Amsterdam: John Benjamins; 2001. p. 137-158.
- Pierrehumbert J. More than 70 years of probabilistic phonology. In: Dresher BE, editor. The Oxford handbook of the history of phonology. Oxford: Oxford University Press; 2022. p. 639-655.
- Pierrehumbert J. Probabilistic phonology: discrimination and robustness. In: Bod R, Hay J, Jannedy S, editores. Probability theory in linguistics. Cambridge: The MIT Press; 2003. p. 177-228.
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria; 2023. Disponível em: <https://www.R-project.org/>.
- Silva AHP. Para a descrição fonético-acústica das líquidas no português brasileiro: dados de um informante paulistano [dissertação]. Campinas: Universidade Estadual de Campinas; 1996.
- Silveira F, Seara IC. Vogal de apoio em grupos consonantais CCV no português brasileiro. Revista da ABRALIN. 2017;7(1). [citado 06 jun. 2023]. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/977>.

Stampe D. A dissertation on Natural Phonology. New York: Garland Press; 1979.

Turk A, Nakai S, Sugahara M. Acoustic segment durations in prosodic research: a practical guide. In: Sudhoff S, et al., editores. Methods in empirical prosody research. Berlin, Boston: De Gruyter; 2006. p. 1-28.

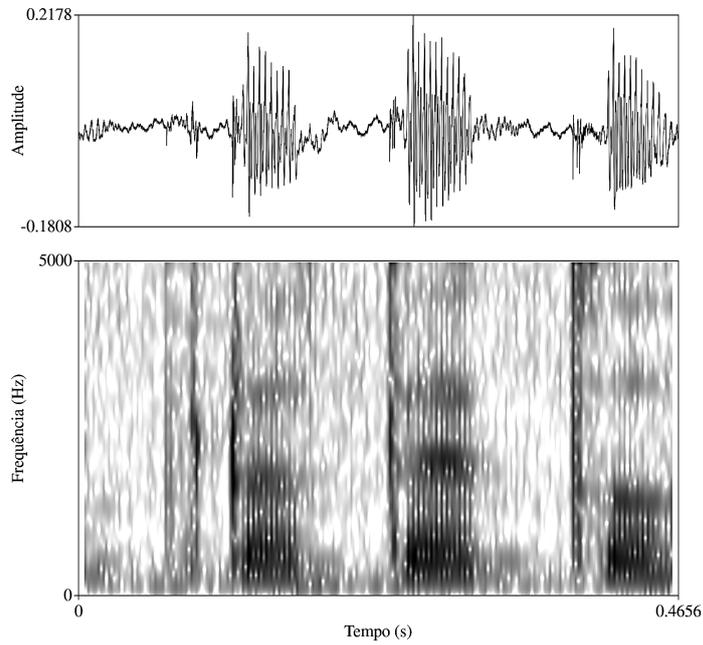
Viaro ME, Guimarães-Filho ZO. Análise quantitativa da frequência dos fonemas e estruturas silábicas portuguesas. Estudos Linguísticos. 2007;26(1):30-36.

## ANEXO



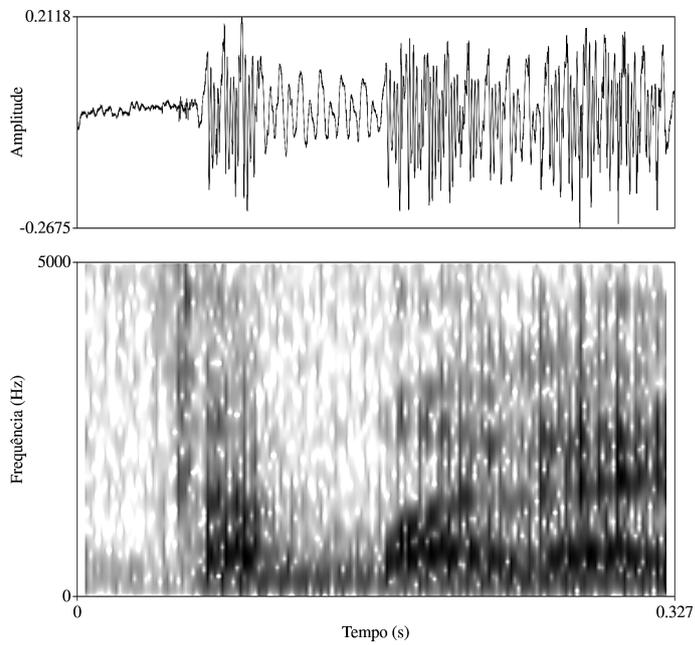
Fonte: Elaborado pelo autor.

Figura A1 - Forma de onda e espectrograma da palavra “preço”, contendo epêntese no EOL.



Fonte: Elaborado pelo autor

Figura A2 - Forma de onda e espectrograma da palavra “protetor”, contendo EOL com padrão CCV.



Fonte: Elaborado pelo autor.

Figura A3 - Forma de onda e espectrograma da palavra “trabalhar”, com apagamento do tepe no EOL.

# O alçamento não assimilatório das vogais /e, o/ pretônicas no português do sul do Brasil: uma abordagem fonológica de contraste e aperfeiçoamento

## *The non-assimilatory raising of pretonic vowels /e, o/ in southern Brazilian Portuguese: a contrast and enhancement approach*

Elisa Battisti\*

Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil

**Resumo:** O artigo trata do alçamento das vogais médias pretônicas /e, o/ não desencadeado por vogal alta seguinte (*senhora* ~ *s[ɨ]nɔra*, *colher* ~ *c[u]lher*) no português de Porto Alegre. A análise baseia-se em Purnell e Raimy (2015), Purnell, Raimy e Salmons (2019), modelo que segue a abordagem de contraste e aperfeiçoamento de Hall (2011) e opera com traços fonológicos privativos, organizados conforme a geometria de traços de Avery e Idsardi (2001). Os objetivos da análise são (i) buscar, na representação fonológica, explicações para as proporções de alçamento observadas, relativamente menores para /e/ do que para /o/, e (ii) representar os contrastes vocálicos do português brasileiro em um modelo com traços privativos pouco explorado na literatura de fonologia do português. A análise mostra que a vogal /o/ é mais suscetível ao alçamento porque é menos marcada do que /e/ no sistema de contrastes da língua. Uma vantagem do modelo usado está em sua estrutura implicacional de traços privativos (gestos), dimensões (contrastos) e níveis superordenados, que explica e restringe a atividade fonológica possível na língua. Uma limitação da análise foi marcar a dimensão Altura da Língua em dois *tiers* da hierarquia contrastiva, para dar conta dos quatro contrastes de altura do português.

**Palavras-chave:** Representação fonológica. Vogais médias /e, o/. Alçamento não assimilatório. Português do sul do Brasil.

**Abstract:** The paper is about the raising of pretonic mid vowels /e, o/ not triggered by a following high vowel (*senhora* ~ *s[ɨ]nɔra* 'lady', *colher* ~ *c[u]lher* 'spoon') in Porto Alegre Portuguese. The analysis is based on Purnell and Raimy (2015), Purnell, Raimy and Salmons (2019), a model that follows the contrast and enhancement approach of Hall (2011) and operates with privative phonological features, organized according to the feature geometry of Avery and Idsardi (2001). The objectives of the analysis are (i) seek, in phonological representation, motivations for the observed frequencies of rising, relatively smaller for /e/ than for /o/, and (ii) represent the vowel contrasts of Brazilian Portuguese in a model with privative features little explored in the Portuguese phonology literature. The analysis shows that /o/ is more susceptible to raising because it is less marked than /e/ in the contrast system. An advantage of the used model is its implicational structure of privative features (gestures), dimensions (contrasts) and superordinate levels, which explains and restricts the phonological activity possible in the language. A limitation of the analysis was marking the Tongue Height dimension twice in the contrastive hierarchy to account for the four contrasts of height of Portuguese.

---

\* Professora Associada, Departamento de Linguística, Filologia e Teoria Literária, Instituto de Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil/CNPq; [elisa.battisti@ufrgs.br](mailto:elisa.battisti@ufrgs.br)

**Keywords:** Phonological representation. Mid vowels /e, o/. Non-assimilatory raising. Portuguese from southern Brazil.

## 1 INTRODUÇÃO

O alçamento não assimilatório das vogais médias /e, o/ em sílaba pretônica ( $p[e]queno \sim p[i]queno$ ,  $g[o]verno \sim g[u]verno$ ), posição em que o contraste entre as médias altas e baixas é neutralizado em português ( $v[\varepsilon]la-v[e]leiro$ ,  $r[\text{ɔ}]cha-r[o]choso$ , cf. Camara Jr., 1970), mereceu nossa atenção em investigação anterior (Battisti; Perozzo; Cunha, 2020). Como outros estudos do alçamento não assimilatório (doravante, apenas *alçamento*) com dados de português de Porto Alegre, comunidade de fala situada no sul do Brasil (Klunck, 2007; Bisol, 2009; Cruz, 2010; Monaretto, 2013; Silva, 2014; Biasibetti, 2014), analisamos dados de produção na linha variacionista laboviana (Labov, 1972). Diferentemente daqueles estudos, que centraram atenção nos efeitos do léxico e da frequência de certas palavras sobre o alçamento, detivemo-nos nas regularidades do ambiente fonético e no papel de certos segmentos adjacentes às vogais afetadas pelo processo. Mostramos que, quando aplicado ao português de Porto Alegre<sup>1</sup>, o alçamento é tanto explicado quanto condicionado por múltiplos fatores, entre eles a coarticulação de sequências CV (Goldstein; Byrd; Saltzman, 2006) e a dispersão vocálica (Maddieson, 1984; Lindblom, 1986). Contudo, não tratamos da diferença na proporção de alçamento entre as vogais, relativamente menor para /e/ do que para /o/, fato atestado em todas as demais análises referidas. Esse é um dos objetivos do presente trabalho: explicar por que /o/ eleva-se mais do que /e/.

A análise aqui realizada, na linha de contraste e aperfeiçoamento de Hall (2011), mostrará que a representação fonológica das vogais médias e a natureza menos marcada de /o/ em relação a /e/ torna aquela vogal mais suscetível à variação do que essa. Está aí outro objetivo do trabalho: representar os contrastes fonológicos do sistema vocálico do português brasileiro (PB) em um modelo na linha de Hall (2011), proposto por Purnell e Raimy (2015), Purnell, Raimy e Salmons (2019) com traços privativos, ainda pouco explorado na literatura de fonologia do português. Assim fazendo, espera-se contribuir para esclarecer as potencialidades e limites do modelo na análise da atividade fonológica envolvendo o vocalismo átono do português.

Sendo o alçamento um processo de variação fonológica, importa esclarecer, de partida, alguns pressupostos teóricos assumidos no artigo. Com Labov (1972, 1981, 1994) e Kiparsky (2016), concebe-se a variação como mecanismo da mudança fonológica que atua conforme a visão neogramática: afeta todas as palavras que incluem a unidade fônica em questão, de acordo com o ambiente fonético, a não ser que restrições estruturais e forças sociais previnam a aplicação do processo. Havendo indícios de difusão gradual da mudança fônica através do léxico, pode-se resolver o paradoxo distinguindo-se mudança fonológica abstrata de mudança por regras de output de nível baixo (*low level rules*, em inglês), que podem, por exemplo, contribuir para sinalizar distinções fonêmicas relevantes em certos contextos. Nesse sentido, e parafraseando o que já afirmamos no trabalho anterior (Battisti; Perozzo; Cunha, 2020), não negamos os efeitos do léxico e da frequência de certas palavras sobre o alçamento. Apenas não atribuímos ao léxico papel central na condução do processo.

<sup>1</sup> As proporções de alçamento não assimilatório são relativamente baixas no português de Porto Alegre, inferiores a 10% para /e/, em torno de 20% para /o/, como atestam os estudos brevemente revisados na seção 2 deste artigo.

Como bem observa Kiparsky (2016), embora processos de redução e enfraquecimento afetem palavras recorrentes mais frequentemente, isso não significa que as palavras mais frequentes liderem e expliquem a mudança, já que tal fato (‘mais recorrente’ igual a ‘mais afetado’) é observado também na variação estável, em que mudança e sua difusão não estão envolvidas, a exemplo do processo abordado no presente artigo.

O modelo de representação fonológica seguido aqui, de Purnell e Raimy (2015), Purnell, Raimy e Salmons (2019), na linha de Hall (2011), considera fonologia um sistema de memória do tipo associativa, uma capacidade de recuperação de conhecimento de acesso direto. O conhecimento fonológico é representado em termos de estruturas relacionais: *tiers* ou camadas, em cadeias hierarquizadas de traços contrastivos, endereçáveis a partir de seu conteúdo. As regras fonológicas operam sobre esse conhecimento, incrementando os contrastes marcados na representação e definindo a substância fonética dos segmentos. Ou seja, tomam-se o fonológico, o fonético-fonológico e o fonético como níveis de representação relacionados, mas distintos: distingue-se o nível de representação fonológica (das oposições contrastivas) do nível fonético-fonológico (das complementações e aperfeiçoamentos), e esses, do nível da implementação fonética propriamente dita (*continuum* de realizações), o que se aproxima de outros modelos de gramática, como o de Bermúdez-Otero (2006, 2013), por exemplo.

## 2 O ALÇAMENTO

Neutralizado o contraste entre as vogais médias altas /e, o/ e baixas /ɛ, ɔ/ na pauta pretônica, observa-se variação, na implementação fonética, entre vogais médias altas e médias baixas (*s[e]tembro* ~ *s[ɛ]tembro*, *c[o]leção* ~ *c[ɔ]leção*), e entre médias altas e vogais alçadas, essas resultantes ou do processo assimilatório de harmonia vocálica (*s[e]rviço* ~ *s[i]rviço*, *c[o]stume* ~ *c[u]stume*), ou de alçamento não assimilatório (*t[e]soura* ~ *t[i]soura*, *t[o]mate* ~ *t[u]mate*). Conforme Silva (2021, p. 30), a “diferença de altura das vogais pré-acentuadas se constitui numa forte marca regional”. A partir dessa diferença, distinguem-se variedades de PB faladas ao sul do Brasil, em que predomina a variação entre médias altas e alçadas, das variedades das demais regiões do país, em que há a tripla possibilidade de realização das médias: “podem-se ouvir, por exemplo, *c[u]légio*, *c[o]légio*, *c[ɔ]légio*, *r[i]vista*, *r[e]vista* ou *r[ɛ]vista*.” (Silva, 2021, p. 30, *italicos da autora*).

O português de Porto Alegre é uma das variedades de PB do sul do Brasil. Apresenta variação entre médias altas e alçadas nas sílabas pretônicas, com frequências de alçamento de moderadas a baixas, sejam elas resultantes de harmonia vocálica ou de alçamento não assimilatório. Não se capta, de oitiva, a tripla possibilidade de realização das médias nessa variedade de português.

O processo de interesse neste artigo, o alçamento não assimilatório, vem sendo investigado no PB de Porto Alegre desde Klunck (2007), tendo recebido diferentes designações: alçamento sem motivação aparente (Klunck, 2007; Bisol 2009; Cruz, 2010; Battisti; Perozzo; Cunha, 2020), elevação sem motivação aparente (Biasibetti, 2014; Cunha, 2018), elevação não desencadeada por vogal alta seguinte (Cunha, 2022), elevação não harmônica (Battisti; Cunha, 2023). Centraremos a revisão da literatura no estudo de Cunha (2022), não só porque confirma o verificado nos demais estudos em termos de condicionamentos linguísticos e sociais do alçamento não assimilatório, mas

também porque realiza análise de produção em tempo real e em tempo aparente<sup>2</sup> (cf. Labov, 1994), utilizando, para a primeira análise, dados coletados nos anos 1990 (VARSUL<sup>3</sup>) e 2015-2019 (LínguaPOA<sup>4</sup>), e, para a segunda, dados do *corpus* mais recente, o que permitiu confirmar o estatuto do alçamento não assimilatório como processo de variação estável no português de Porto Alegre.

## 2.1 O alçamento não assimilatório variável no português de Porto Alegre

Como afirmamos na introdução, estudos sobre o alçamento não assimilatório no português de Porto Alegre (Klunck, 2007; Bisol, 2009; Cruz, 2010; Monaretto, 2013; Biasibetti, 2014) atestam efeitos de frequência e tratam o processo como difusão lexical, ao mesmo tempo em que comprovam o condicionamento fonético do processo por consoantes palatais, velares, labiais, precedentes ou seguintes. O estudo de Cunha (2022) confirma esses resultados.

Para a análise em tempo real, a autora extraiu, de oitiva, 1075 contextos de alçamento de /e/ e 1111 contextos de alçamento de /o/ de 12 entrevistas do VARSUL e 1513 contextos de alçamento de /e/ e 1448 contextos de alçamento de /o/ de 12 entrevistas do LínguaPOA, distribuídas equilibradamente em dois sexos/gêneros (masculino, feminino) e três grupos etários (20-39 anos, 40-59 anos, 60 anos ou mais). Controlou as variáveis previsoras linguísticas *altura da vogal seguinte*, *contexto precedente*, *contexto seguinte*, *nasalidade*, *tipo de sílaba*, *distância da sílaba tônica*, *posição da sílaba*, e as variáveis previsoras sociais *gênero*, *escolaridade*, *faixa etária* e *zona*. Realizou análise de regressão logística de efeitos mistos na plataforma R (R Core Team, 2021), com *informante* e *palavra* como variáveis aleatórias. As proporções totais de aplicação estão na Tabela 1.

Tabela 1 - Proporções totais de alçamento não assimilatório na análise em tempo real de Cunha (2022).

/e/		/o/	
VARSUL (1990)	LínguaPOA (2015-2019)	VARSUL (1990)	LínguaPOA (2015-2019)
9,5%	5,8%	23,5%	22,5%

Fonte: Elaborada pela autora.

<sup>2</sup> Conforme Labov (1994), na análise em tempo real, estudo de tendência, examinam-se dados de fala de informantes de mesmo perfil social (sexo/gênero, idade, escolaridade etc.), coletados na mesma comunidade, mas em pelo menos dois períodos (ou sincronias) distintos. Na análise em tempo aparente, os dados examinados são de um único período (uma só sincronia). Compara-se a realização da variável de interesse por diferentes grupos etários, no pressuposto de que o padrão de uso dos mais jovens representa o que a língua poderá ser no futuro.

<sup>3</sup> VARSUL (<https://www.varsul.org.br/>) é um banco de dados de fala de informantes de ambos os sexos (masculino, feminino), com diferentes níveis de escolaridade (Fundamental I, Fundamental II, Médio) e idades (25 a 50 anos, 50 anos ou mais), residindo nas capitais e algumas cidades do interior dos estados do Rio Grande do Sul, Santa Catarina e Paraná. As entrevistas da amostra base do VARSUL foram realizadas no início dos anos 1990.

<sup>4</sup> LínguaPOA (<https://www.ufrgs.br/linguapoa/>) é um acervo de entrevistas sociolinguísticas de informantes de Porto Alegre realizadas entre 2015 e 2019, estratificadas em 4 zonas (centro, norte, sul, leste), 2 bairros por zona (por renda média mensal em salários mínimos), 3 grupos etários (20-39 anos, 40-59 anos, 60 ou mais anos), 3 níveis de escolaridade (fundamental, médio, superior), 2 gêneros (masculino e feminino). As atuais 103 entrevistas atendem a todos os critérios de estratificação nos níveis médio e superior de escolaridade.

A análise em tempo real de Cunha (2022) mostrou diminuição na proporção total de alçamento não assimilatório de /e/ entre os anos 1990 e 2015-2019 e taxas muito similares de aplicação do processo à vogal /o/ no português de Porto Alegre nesses dois períodos. A análise em tempo aparente, também efetuada pela autora, confirma essas proporções.

Para a análise em tempo aparente, Cunha (2022) extraiu de oitiva 3751 contextos de alçamento de /e/ e 3292 contextos de alçamento de /o/ de 32 entrevistas sociolinguísticas do LínguaPOA (2015-2019). A análise de regressão logística de efeitos mistos contou com as mesmas variáveis aleatórias e as mesmas variáveis predictoras linguísticas da análise em tempo real, mas com diferentes variáveis predictoras sociais: *sexo/gênero* (masculino, feminino), *escolaridade* (médio/fundamental e superior), *faixa etária* (20-39 anos, 60 anos ou +) e *zona* (Central, Norte, Sul e Leste). As proporções totais de alçamento foram de 7% para /e/, correlacionado a *contexto precedente* e *contexto seguinte*, e de 21% para /o/, correlacionado a *contexto precedente* e *contexto seguinte*. Os Quadros 1 e 2, elaborados pela autora, sintetizam os resultados das análises de regressão por ela efetuadas.

Quadro 1 – Síntese dos resultados de Cunha (2022) para a vogal /e/.

Vogal /e/			
	<i>Análise em tempo aparente</i>	<i>Análise em tempo real (VARSUL)</i>	<i>Análise em tempo real (LínguaPOA)</i>
<b>Contexto precedente</b>	líquidas e nasais desfavorecem; oclusivas favorecem	oclusivas favorecem	oclusivas favorecem
<b>Contexto seguinte</b>	líquidas desfavorecem; vogais favorecem	vogais favorecem	vogais favorecem
<b>Faixa etária</b>	-	adultos (40 a 59 anos) desfavorecem	-

Fonte: Cunha (2022, p. 59).

Quadro 2 – Síntese dos resultados de Cunha (2022) para a vogal /o/.

Vogal /o/			
	<i>Análise em tempo aparente</i>	<i>Análise em tempo real (VARSUL)</i>	<i>Análise em tempo real (LínguaPOA)</i>
<b>Altura da vogal seguinte</b>	-	médias favorecem	-
<b>Contexto precedente</b>	soantes e vogais desfavorecem	lateral desfavorece	-
<b>Contexto seguinte</b>	laterais desfavorecem; vogais favorecem	-	vogais favorecem
<b>Gênero</b>	-	-	masculino favorece
<b>Faixa etária</b>	-	adultos (40 a 59 anos) desfavorecem	adultos (40 a 59 anos) masculinos favorecem

Fonte: Cunha (2022, p. 59).

Chama atenção, nesses resultados, o fato de não haver correlação do alçamento de ambas as vogais com variáveis predictoras sociais na análise em tempo aparente, apenas na análise em tempo real<sup>5</sup>, o que reforça a ideia de estabilização do processo na

<sup>5</sup> As análises estatísticas realizadas por Cunha (2022) ocorreram em duas etapas: (1) teste de qui-quadrado (de Pearson), com todas as variáveis, e (2) análise multivariada de regressão logística, em modelos lineares de efeitos mistos (função glmer). Na etapa 2, foram incluídas nos modelos de regressão apenas as variáveis que apresentaram valor-p significativo – igual ou menor do que 0,05 –

comunidade. Em relação às variáveis predictoras linguísticas, o processo correlaciona-se aos contextos fonológicos precedente e seguinte nos dois tipos de análise. Ou seja, além de as proporções totais terem permanecido relativamente estáveis em cerca de 30 anos, atesta-se também a manutenção dos efeitos dos segmentos precedentes e seguintes sobre o alicamento e a ausência de condicionamento social na amostra mais recente. A sistematicidade expressa por esses resultados sugere a possibilidade de concebê-lo como resultante de processo fonético-fonológico no português de Porto Alegre. A regularidade de sua aplicação pode estar relacionada ao próprio sistema de contrastes do português e à neutralização das vogais médias em sílabas pretônicas. Antes de testar essa hipótese, é relevante resgatar alguns achados de análises fonológicas do vocalismo átono do PB, que orientarão a análise aqui realizada.

## 2.2 O alicamento variável e a fonologia das vogais do português

Camara Jr. (1970) propõe haver sete contrastes vocálicos em sílabas tônicas /i, e, ε, a, ɔ, o, u/ em português, reduzidos a cinco, /i, e, a, o, u/, em posição pretônica, pela neutralização do contraste entre as vogais médias. Essa hipótese está na base de análises fonológicas como a de Lee e Oliveira (2003), sobre a variação interdialetoal e intradialetoal na realização das médias /e, o/ pretônicas. Para os autores:

Costuma-se dizer que os dialetos do PB podem ser divididos em dois grandes grupos, o do norte e o do centro-sul [...] os dialetos do norte se caracterizariam pelas vogais pretônicas médias de timbre aberto [...] enquanto que os dialetos do centro-sul se caracterizariam por apresentarem, para essas mesmas vogais, um timbre fechado [...] A realidade é que nos dois grandes grupos dialetais podemos ter [ɔ~o~u] e [ε~e~i], em posição pretônica. (Lee; Oliveira, 2003, p. 67-68).

Os autores analisam, pela Teoria da Otimidade (TO) (Prince; Smolensky, 1993; McCarthy; Prince, 1995), os processos de harmonia vocálica (HV) e redução vocálica ('alçamento não assimilatório', neste artigo), como também abaixamento. Os traços fonológicos referidos pelas restrições que, na proposta dos autores, interagem na gramática constam no Quadro 3 (as siglas BK, HI, LO, RD correspondem aos traços bivalentes [*±posterior*], [*±alto*], [*±baixo*], [*±arredondado*], respectivamente).

---

na etapa 1. Por isso, por exemplo, na análise em tempo aparente dos dados de /e/, apenas as variáveis *altura da vogal seguinte*, *contexto precedente*, *contexto seguinte* e *tipo de sílaba* foram incluídas nos modelos; já na análise em tempo aparente dos dados de /o/, as variáveis *altura da vogal seguinte*, *contexto precedente*, *contexto seguinte*, *tipo de sílaba*, *distância da sílaba tônica*, *posição da sílaba*, *gênero* e *escolaridade* foram incluídas nos modelos. Os modelos estatísticos de regressão de Cunha (2022), com os resultados quantitativos (frequências de alicamento por nível de cada variável predictor, mais as estimativas) estão disponíveis em:

<http://www.bibliotecadigital.ufrgs.br/da.php?nrb=001145032&loc=2022&l=42758e33389af39e>.

Acesso em 13 jun. 2024.

Quadro 3 – Inventário do sistema vocálico do PB conforme Lee e Oliveira (2003).

		[-BK]	[+BK]	
[+HI]	[+ATR]	i	u	[-LO]
[-HI]	[+ATR]	e	o	
	[-ATR]	ɛ	ɔ	
	[-ATR]		a	[+LO]
		[-RD]	[+RD]	

Fonte: Lee e Oliveira (2003, p. 74).

Na análise dos autores, as vogais de input são plenamente especificadas. No que interessa mais diretamente ao presente artigo, vale destacar que, na linha de análise seguida por Lee e Oliveira (2003), as formas de output com vogais médias alçadas são candidatos com mudança na especificação do traço  $[\pm\text{alto}]$ , de [-HI] para [+HI]. Os candidatos são selecionados por satisfazerem restrições que proíbem vogal média [-HI, -LO] em sílaba pretônica, quando essas estão altamente hierarquizadas.

Para os autores, teorias fonológicas como a TO, voltadas ao output e centradas na interação das restrições no processamento (paralelo) necessário ao mapeamento input-output, conseguiriam dar conta da variação interdialeto, gerada por múltiplas gramáticas: cada comunidade de fala seria relativamente homogênea, teria a sua gramática (hierarquia de restrições) particular, e a variação resultaria de “alterações nas hierarquias entre as restrições ou através da utilização de conjuntos diferentes de restrições” (Lee; Oliveira, 2003, p. 88-89) de uma comunidade para outra. Contudo, apontam a incapacidade do modelo de lidar com a variação intradialeto. As comunidades de fala não são homogêneas. Seria preciso, em uma concepção de gramática como a da TO, lidar com múltiplas gramáticas em uma única comunidade, o que requereria esclarecer até que ponto ajustes nas hierarquias de restrições seriam possíveis, limite teórico de que o modelo, na versão (Prince; Smolensky, 1993; McCarthy; Prince, 1995) usada pelos autores, não dispõe.

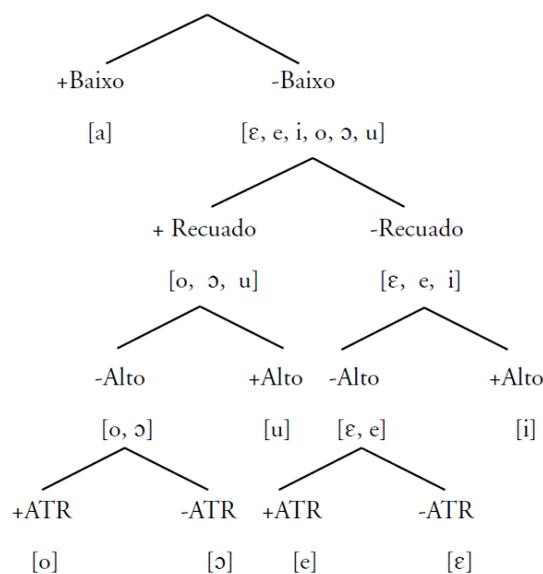
Outro fato relevante, não mencionado pelos autores, mas digno de registro é o seguinte: os padrões quantitativos de realização das variantes resultam não apenas de restrições estruturais, mas também de forças sociais relevantes. O processamento cognitivo da variação fonológica, em termos de produção e percepção, encontra na fonologia das línguas naturais apenas parte de suas motivações<sup>6</sup>. No que diz respeito a nosso objeto de interesse, o alçamento não assimilatório das vogais médias pretônicas /e, o/, os aspectos relevantes parecem ser, de um lado, os contrastes vocálicos e sua representação por traços, neutralizados na pauta pretônica; de outro, um modelo de gramática fonológica que separe o nível do sistema de contrastes do nível relacionado à atividade fonológica da língua (das regras ou processos fonológicos). Além disso, a percepção/produção de contrastes fonológicos e a percepção/aplicação de regras fonológicas, categóricas ou não, embora intimamente relacionadas, têm estatuto particular, evidenciado no próprio processo de aquisição da linguagem, bastante dependente dos dados linguísticos a que a criança é exposta.

Sobre a aquisição das vogais pretônicas do PB, Matzenauer e Miranda (2007) afirmam que essa se dá em três estágios. A partir do exame de dados de crianças do Rio Grande do Sul, expostas a uma variedade de PB em que vogais médias abertas

<sup>6</sup> De acordo com Coetzee (2016), fatores linguísticos (gramaticais) dirigem a variação e fatores não linguísticos (sociais) têm efeitos sobre os processos, posição que se assume aqui.

ocorrem apenas em sílaba tônica, as autoras propõem que, no primeiro estágio, as crianças adquirem o contraste entre /a/, /i/, /u/; no segundo, entre /a/, /i/, /u/, /e/, /o/, no terceiro, o sistema completo: /a/, /i/, /u/, /e/, /o/, /ɛ/, /ɔ/. Ou seja, o contraste entre as vogais médias é adquirido mais tardiamente. Na série das médias, as médias baixas são as últimas adquiridas. Segundo Matzenauer (2009), os traços [alto] e [baixo] funcionam como monovalentes na representação dos contrastes no primeiro estágio de aquisição. A aquisição das médias altas no segundo estágio dá-se pela ativação do valor binário dos traços [±alto], [±baixo]. O subsistema das pretônicas estabiliza-se mais tardiamente do que o das tônicas, pela variabilidade na realização dos segmentos. Segundo a autora, as crianças são sensíveis a processos fonológicos envolvendo as vogais pretônicas (HV, elevação, neutralização), aplicados conforme os estágios de aquisição e condicionados pela sílaba, pelo pé métrico e pela estrutura interna do segmento. Matzenauer (2009) explica que as vogais pretônicas são mais sensíveis à aplicação dos processos “porque ficam fora do pé portador do acento primário da palavra prosódica” (Matzenauer, 2009, p. 107).

Lee (2008, 2012) considera as três etapas de aquisição das pretônicas propostas por Matzenauer e Miranda (2007) e Matzenauer (2009) para hierarquizar os traços contrastivos no sistema vocálico do PB e, a partir do sistema assim representado, dar conta dos processos que afetam as vogais nessa língua. Entre outras generalizações sobre o comportamento das classes de vogais em processos fonológicos no PB, Lee (2012) afirma que as vogais médias nas posições pretônica e tônica costumam ser alvos de processos fonológicos; as vogais médias e as vogais altas são gatilhos em tais processos; a vogal baixa é sempre neutra nos processos fonológicos; as vogais altas podem apenas desencadear a HV e não sofrem alternâncias no PB. Levando em conta essas generalizações e as três etapas de aquisição, Lee (2008, 2012) propõe que, na hierarquia dos contrastes vocálicos do PB, as vogais médias são marcadas para todos os traços no inventário da língua, e a vogal baixa é a menos marcada. A hierarquia proposta pelo autor está na Figura 1, com os traços bivalentes [±baixo] > [±recuado] > [±alto] > [±ATR].



Fonte: Lee (2012, p. 10).

Figura 1 – Hierarquia contrastiva dos traços das vogais do PB (Lee, 2008).

O autor baseia-se nessa hierarquia para, como antes fizeram Lee e Oliveira (2003), dar conta da variação interdialetoal das médias em PB pela TO, via atuação, na gramática, de restrições voltadas a essas vogais, as mais marcadas no sistema de contrastes.

Já Wetzels (2011), ao tratar da representação da altura vocálica em PB, examina a funcionalidade global do contraste entre as médias em verbos e não verbos. Sobre o contraste entre médias altas e baixas, o autor afirma que um sistema com traços bivalentes como o  $[\pm\text{ATR}]$  seria insuficiente para representar especialmente a relação entre formas verbais com alternância entre três realizações na vogal da raiz, como *nós s[e]rvimos, ele s[ɛ]rve, eu s[i]rvo*. Essa é uma das razões para o autor valer-se do sistema de traços de Clements (1991) (Figura 2), com graus de abertura escalares, em vez de um traço bivalente como  $[\pm\text{ATR}]$ , para representar os contrastes de altura das vogais. Cada grau de abertura, por seu turno, é marcado com os traços bivalentes  $[\pm\text{aberto1}]$ ,  $[\pm\text{aberto2}]$ ,  $[\pm\text{aberto3}]$ .

	i/u	e/o	ɛ/ɔ	a
aberto <sub>1</sub>	-	-	-	+
aberto <sub>2</sub>	-	+	+	+
aberto <sub>3</sub>	-	-	+	+

Fonte: Adaptada de Wetzels (2011, p. 353).

Figura 2 – Sistema vocálico do PB em sílaba tônica.

Wetzels (2011) propõe que a neutralização seja concebida como substituição, não desassociação de traço: os valores dos traços contrastivos são substituídos pelos valores opostos no *tier* em que o contraste relevante é definido. Disso resultam três subsistemas vocálicos plenamente especificados, um para a neutralização em sílaba acentuada, que ocorre em verbos, e dois para sílaba não acentuada em não verbos (neutralização em sílabas átonas em final de palavra e neutralização nas demais posições átonas da palavra).

A proposta do autor atende à funcionalidade da neutralização vocálica em PB, mas herda, dos sistemas de traços binários, a previsão de que a atividade fonológica seja possível nos dois lados no contraste, o positivo e o negativo, o que implica operar com algum grau de estipulação nas análises. Um sistema de traços privativos, como o que será usado neste artigo, tem a vantagem de possibilitar marcar, em um dado *tier* contrastivo, apenas o traço fonologicamente ativo na língua, conforme Purnell, Raimy e Salmons (2019), fortalecendo a conexão entre traços, a marcação de contrastes e a atividade fonológica.

Das análises fonológicas do vocalismo átono do PB aqui brevemente revisadas, destacam-se os seguintes achados: (i) a altura é a propriedade relevante ao contraste e à implementação fonética das vogais médias; (ii) a altura é marcada tardiamente no sistema de contrastes ao longo da aquisição da linguagem; (iii) a marcação tardia sugere a plausibilidade de operar com vogais mais e menos marcadas no sistema de contrastes;

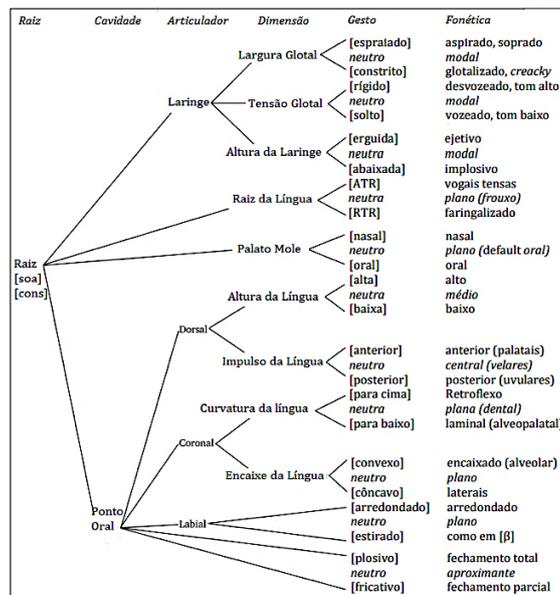
(iv) a altura das vogais médias é variavelmente alterada por processos fonético-fonológicos, a partir do sistema de contrastes; e (v) a representação dos contrastes de altura vocálica por traços binários desafia as análises no estabelecimento de conexões entre a atividade fonológica e a marcação de contrastes no sistema da língua. A questão que se apresenta, então, para a análise aqui empreendida é: como expressar, de um lado, a relação entre os contrastes segmentais e a variabilidade de superfície, e, de outro, a maior susceptibilidade de /o/ do que /e/ a processos como a elevação?

Assume-se, nesta análise, que a representação motiva e limita as flutuações observadas, distinguindo o fonológico (contrastes), o fonético-fonológico (processos) e o fonético (*continuum* fônico, gradiente). Esse é o pressuposto de Purnell e Raimy (2015) e Purnell, Raimy e Salmons (2019), modelo que opera com traços privativos, como se esclarece a seguir.

### 3 MODELO DE ANÁLISE

O modelo de análise aqui seguido, de Purnell e Raimy (2015) e Purnell, Raimy e Salmons (2019), adere à abordagem de contraste e aperfeiçoamento de Hall (2011), para quem os segmentos, representados fonologicamente por traços, realizam-se foneticamente dentro dos limites estabelecidos pelos traços. Ou seja, a realização fonética dos segmentos, que inclui a variação, relaciona-se ao sistema de contrastes da língua: resulta da implementação e aperfeiçoamento dos contrastes, processos fonético-fonológicos influenciados por diferentes fatores.

Purnell e Raimy (2015) e Purnell, Raimy e Salmons (2019) dão corpo à abordagem de Hall (2011) com traços privativos, organizados hierarquicamente na estrutura interna dos segmentos conforme a geometria de traços de Avery e Idsardi (2001). Nessa geometria, além do nível da Raiz e da Cavidade, há três níveis de estruturação segmental: o nível do Articulador, o nível da Dimensão e o nível do Gesto (Figura 3).



Fonte: Adaptada de Purnell e Raimy (2015, p.526) por Battisti e Soares (2022, p. 9).

Figura 3 – Organização segmental com dimensões, gestos e manifestação fonética, de Purnell e Raimy (2015), Purnell, Raimy e Salmons (2019), desenvolvida a partir de Avery e Idsardi (2001).

Laringe é um articulador ligado diretamente à Raiz. Já os outros três articuladores, Dorsal, Labial e Coronal, agrupam-se no nó (de Cavidade) Ponto Oral, que se liga à raiz. O articulador Laringe tem três dependentes, as dimensões Largura Glotal, Tensão Glotal e Altura da Laringe. As dimensões Raiz da Língua e Palato Mole ligam-se diretamente à raiz. As dimensões Altura da Língua e Impulso da Língua, ao articulador Dorsal, as dimensões Curvatura da Língua e Encaixe da Língua, ao articulador Coronal. O articulador Labial não tem dimensões dependentes.

O modelo de Purnell e Raimy (2015) e Purnell, Raimy e Salmons (2019) segue a ideia de Avery e Idardi (2001) de que os contrastes ocorrem no nível da dimensão. Cada dimensão refere um par de grupos musculares antagonistas, por isso apenas um músculo dentro de uma dimensão está ativo em qualquer fone. Os gestos, por seu turno, são instruções motoras aos articuladores. Representam a ação de um único grupo muscular e são os últimos constituintes das representações fonológicas. Os gestos correspondem a traços em outros modelos, mas não são binários e não expressam contrastes fonológicos. São adicionados à representação no componente fonético-fonológico, ou por operações *default*, ou por regras fonológicas.

As categorias superiores Raiz, Cavidade e Articulador, bem como Dimensão, não são pronunciáveis. As dimensões precisam ser completadas com gestos para a sua implementação fonética. De acordo com Purnell e Raimy (2015), a completação ocorre no nível fonético-fonológico de representação. Corresponde à atribuição de estrutura ao segmento por processos adicionais. As dimensões são completadas com um de dois gestos possíveis, ou deixadas inertes (neutras). No nível fonético, esses gestos convertem-se em unidades contínuas no sinal da fala. “Cada gesto tem seu resultado articulatorio ou acústico e também há uma opção ‘neutra’, quando nenhum dos músculos em um par está inervado. A opção fonética ‘neutra’ é equivalente a ser fonologicamente não marcado no sentido estrutural.” (Purnell; Raimy, 2015, p. 7, aspas simples dos autores).

As regras fonológicas que completam as dimensões com gestos são específicas de língua e sensíveis ao contexto. Convertem a marcação contrastiva, que é uma representação fonologicamente esparsa, armazenada na memória, em representações fonéticas mais completamente especificadas, que funcionam como instruções aos sistemas de controle motor. Ativar um certo músculo de qualquer dimensão, e não o membro antagonístico do par, é parte da aquisição do sistema fônico de qualquer língua.

Outros gestos, referentes a dimensões não contrastivas, podem ser adicionados no nível fonético-fonológico por meio de processos de aprimoramento (*enhancement*, em inglês), que aumentam a saliência perceptual dos contrastes fonêmicos e facilitam a pronunciabilidade de categorias não especificadas. Por exemplo, em uma língua que contenha /l/ e /r/ no conjunto de soantes contrastivas, a dimensão Encaixe da Língua (completada com o gesto [côncavo]) representa /l/, distinguindo esse segmento de /r/, não especificado na dimensão Encaixe da Língua. Subespecificado no nível da dimensão, /r/ fica mais suscetível à variabilidade, conforme Natvig (2020).

Purnell e Raimy (2015) e Purnell, Raimy e Salmons (2019) assumem que a marcação dos contrastes nas dimensões é feita pelo *Successive Division Algorithm* – SDA (‘Algoritmo de Divisão Sucessiva’, de Dresher (2009)). O SDA é um algoritmo de aquisição de linguagem. Ele estabelece os contrastes fonológicos da língua distinguindo, sucessivamente, os segmentos uns dos outros, com base em um traço

distintivo por vez. Assim fazendo, o SDA opõe marcado<sup>7</sup> a não marcado em cada nível e reduz os fonemas incluídos em um determinado conjunto, a cada iteração. Como consequência e até o estabelecimento de todos os contrastes, haverá subconjuntos de segmentos ainda sem especificação em certos níveis. Outros permanecerão subespecificados até o fim das iterações. Purnell, Raimy e Salmons (2019) resolvem a questão da falta de especificação em certos níveis com o que chamam de ‘marcação nula superordenada’: os membros não especificados para certo contraste em uma dada iteração são representados por seu nó dominante (superordenado).

A primeira oposição contrastiva marcada pelo SDA é estabelecida, conforme Purnell, Raimy e Salmons (2019), entre as classes principais de segmentos. Usam-se os traços [consoante] e [soante] para tanto, não designados para marcação superordenada porque ocorrem como traços de Raiz. Nos níveis sucessivos a partir daí, a marcação superordenada representa o membro não marcado em pares contrastivos, que podem opor um segmento a um conjunto de segmentos ao longo das iterações, ou um segmento a outro segmento, quando as iterações chegam a seu final.

Se o sistema de contrastes não previr um nó articulador primário, categórico e exclusivo para um dado segmento, a estrutura não exige que o segmento tenha um ponto de articulação primário específico. Nesse caso, as estruturas superordenadas na representação servirão como entrada para processos fonético-fonológicos de que poderão resultar a variante ou variantes de certos fonemas.

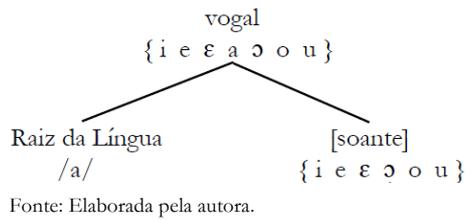
#### 4 ANÁLISE

A análise inicia-se com o estabelecimento dos sete contrastes vocálicos observados em sílabas tônicas no PB, já que o alçamento de /e, o/ em sílabas pretônicas ocorre em ambiente no qual há neutralização do contraste na série das vogais médias.

O estabelecimento dos contrastes leva em conta a ordem de aquisição das vogais do PB conforme Matzenauer e Miranda (2007) e Matzenauer (2009) e também seguida por Lee (2008, 2012). A marcação nas dimensões Raiz da Língua, Impulso da Língua e Altura da Língua será a responsável pelos contrastes entre os sete segmentos vocálicos. Segue-se Purnell, Raimy e Salmons (2019, p. 464) e toma-se a marcação Raiz da Língua como equivalente ao traço (privativo) [RTR], ou [+baixo], na perspectiva gerativista clássica (Chomsky; Halle 1968), Impulso da Língua, como equivalente a [anterior] (i.e. [-posterior]) e Altura da Língua, como [alto] (i.e. [+alto]).

Altura é a dimensão envolvida na marcação do primeiro contraste no sistema. Essa marcação contrasta /a/ com as demais vogais e ocorre na dimensão Raiz da Língua na primeira iteração (Figura 4).

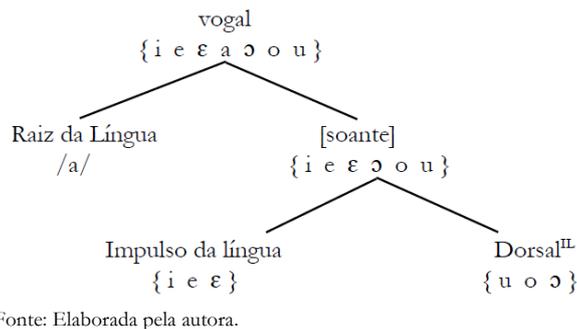
<sup>7</sup> O termo *marcado* é usado, no modelo, para indicar “que um traço distintivo é memorizado para criar contraste. Marcar um contraste constrói estrutura fonológica e fornece alguma, mas não toda a substância fonética ao segmento”. (Purnell; Raimy; Salmons, 2019, p. 459).



Fonte: Elaborada pela autora.  
 Figura 4 – Marcação do contraste de /a/ com as demais vogais do PB (1ª iteração).

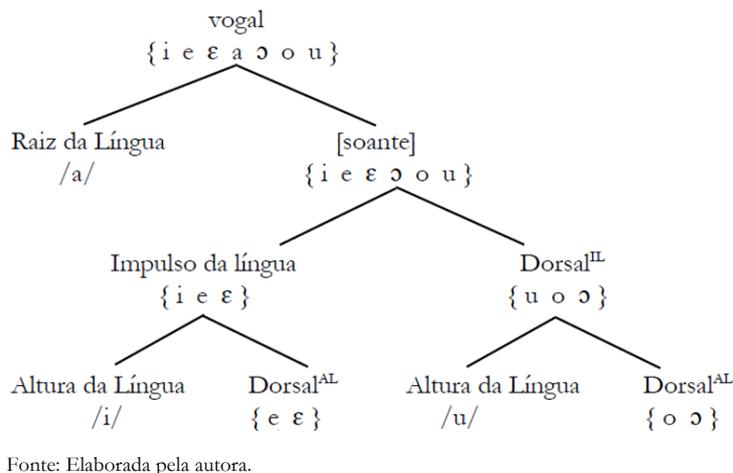
Na Figura 4, o termo ‘vogal’ corresponde à Raiz, por se tratar de uma raiz vocálica. Abaixo de vogal, está o primeiro nível de marcação. O lado marcado (especificado) é o esquerdo. O lado nulo (ainda não especificado) é o direito. O lado nulo é marcado [soante] pela marcação nula superordenada, uma vez que, conforme afirmam Purnell, Raimy e Salmons (2019, p. 465), o traço [soante] “domina [a dimensão] Raiz da Língua na hierarquia (dado que vogais não são [consonantal])”.

Na iteração seguinte, as vogais {i e ε} são marcadas na dimensão Impulso da Língua do articulador Dorsal (Figura 5). O lado nulo é marcado Dorsal<sup>II</sup> pela marcação nula superordenada, significando que o articulador Dorsal é subespecificado com respeito à dimensão Impulso da Língua (II).



Fonte: Elaborada pela autora.  
 Figura 5 – Marcação contrastiva de {i e ε} na dimensão Impulso da Língua (2ª iteração).

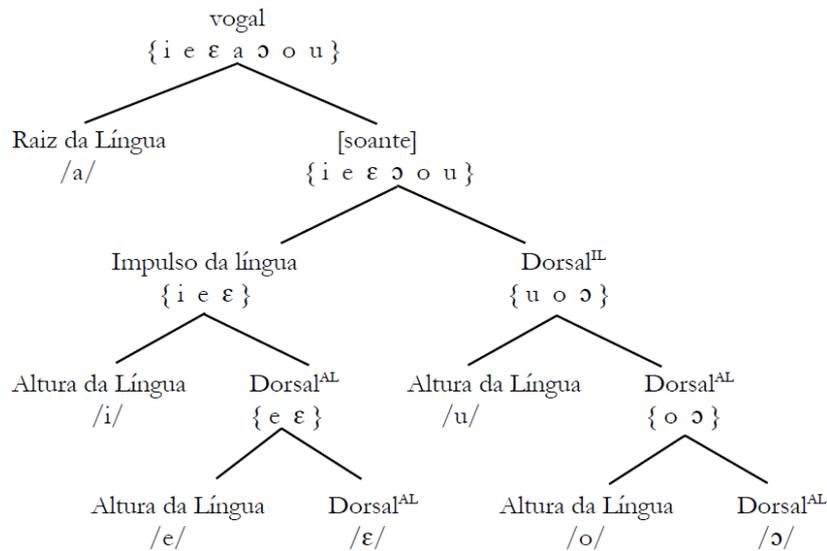
Na próxima iteração, as vogais /i/e /u/ são marcadas na dimensão Altura da Língua do articulador Dorsal (Figura 6).



Fonte: Elaborada pela autora.  
 Figura 6 – Marcação do contraste de /i/ e /u/ na dimensão Altura da Língua (3ª iteração).

Com a marcação das vogais altas /i, u/ na dimensão Altura da Língua na terceira iteração, as vogais médias, situadas no lado nulo da representação (o direito), são marcadas Dorsal<sup>AL</sup> (subespecificadas no articulador Dorsal com respeito à dimensão Altura da Língua (AL)).

Por fim, /e/ e /o/ são marcadas na dimensão Altura da Língua na última iteração (Figura 7).



Fonte: Elaborada pela autora.

Figura 7 – Marcação contrastiva de /e/ e /o/ na dimensão Raiz da Língua (4ª iteração).

As vogais médias altas são marcadas na dimensão Altura da Língua<sup>8</sup> na quarta iteração, no lado esquerdo da representação. As vogais médias baixas situam-se no lado nulo (direito). Por essa razão, são marcadas Dorsal<sup>AL</sup> pela marcação nula superordenada.

Embora a marcação seja igual para os pares de médias /e, o/ e /ε, ɔ/ na última iteração, a cadeia de marcação das vogais na hierarquia contrastiva é diferente. Purnell, Raimy e Salmons (2019) seguem Spahr (2016) e entendem que os traços distintivos na geometria de cada segmento organizam-se em uma cadeia de traços (*feature chain*, em inglês). A ordem de marcação dos traços na hierarquia contrastiva corresponde à ordem dos traços na cadeia. A quantidade de níveis e o que se marca em cada um deles é o que distingue as vogais umas das outras. Os autores implementam a proposta de Spahr (2016) usando números para indicar a sequência de marcação dos traços.

<sup>8</sup> Agradeço a um dos pareceristas anônimos apontar Altura da Língua como a dimensão contrastiva possível em um nível subordinado a um nó de Ponto. A análise com a repetição de Altura da Língua na última iteração pode justificar-se pelos processos fonológicos que afetam a altura das vogais médias em posições átonas e são discutidos na literatura, como faz Wetzels (2011), revisado no presente artigo.

Quadro 4 – Representação de traços das vogais do PB em cadeias numeradas.

/a/	/i/	/u/	/e/	/o/	/ɛ/	/ɔ/
1.vogal	1.vogal	1.vogal	1.vogal	1.vogal	1.vogal	1.vogal
2.Raiz.da.	2.[soante]	2.[soante]	2.[soante]	2.[soante]	2.[soante]	2.[soante
L.	3.Impul.da.	3.Dorsal <sup>IL</sup>	3.Impulso.da.	3.Dorsal <sup>IL</sup>	3.Impulso.da.	]
	L.	4.Altura.da.	L.	4.Dorsal <sup>AL</sup>	L.	3.Dorsal <sup>I</sup>
	4.Altura.da.	L.	4.Dorsal <sup>AL</sup>	5.Altura.da	4.Dorsal <sup>AL</sup>	L
	L.		5.Altura.da.L	.L	5.Dorsal <sup>AL</sup>	4.Dorsal <sup>AL</sup>
						5.Dorsal <sup>AL</sup>

Fonte: Elaborado pela autora.

O Quadro 4 mostra que as cadeias de traços das vogais médias /e, o, ɛ, ɔ/ são mais longas do que as das vogais /a, i, u/. Vale dizer que as vogais médias são relativamente mais marcadas do que as demais vogais.

Comparando-se as cadeias de traços das vogais médias /e, ɛ/ entre si, a vogal média baixa /ɛ/ apresenta mais marcação nula superordenada (i.e., subespecificação) do que a vogal média alta /e/: a vogal /ɛ/ é subespecificada nos *tiers* 2, 4 e 5, a vogal /e/, nos *tiers* 2 e 4. Também na comparação das cadeias de traços das vogais médias posteriores /o, ɔ/ entre si, a vogal média baixa /ɔ/ apresenta mais marcação nula superordenada do que a vogal média alta /o/: a vogal /ɔ/ é subespecificada em todos os *tiers* abaixo do 1, a vogal /o/, nos *tiers* 2, 3 e 4. Assim sendo, pode-se afirmar que ambas as vogais posteriores /o, ɔ/ são menos marcadas do que /e, ɛ/.

Em relação apenas às vogais médias altas /e, o/, a comparação das cadeias de traços mostra que /o/ é mais subespecificada do que /e/: a vogal /e/, como afirmamos, tem marcação nula superordenada nos *tiers* 2 e 4, a vogal /o/, nos *tiers* 2, 3 e 4. Essa diferença no grau de marcação nula de /o/ em relação a /e/ pode explicar por quê /o/ sofre mais alçamento do que /e/. De acordo com Purnell, Raimy e Salmons (2019), segmentos mais marcados têm mais substância fonética e, por isso, apresentam menos variação na implementação. Já segmentos subespecificados “não terão tanta substância fonética, de modo que podem ser mais variáveis na implementação.” (Purnell; Raimy; Salmons, 2019, p. 459). Ou seja, essa diferença na marcação de /e/ em relação a /o/ responde a uma das perguntas deste artigo: a vogal /o/ é mais suscetível ao alçamento variável porque é menos marcada (mais subespecificada) do que /e/.

A representação de traços das vogais do PB em cadeias numeradas (Quadro 4) mostra, portanto, que o contraste de altura entre as vogais médias é marcado no *tier* 5: as médias altas /e, o/ são marcadas na dimensão Altura da Língua e as médias baixas /ɛ, ɔ/ apresentam marcação nula superordenada Dorsal<sup>AL</sup> nessa camada da representação. Havendo neutralização do contraste entre as vogais médias altas e baixas em sílabas átonas, pode-se concebê-la como desligamento ou desassociação dos traços marcados no *tier* 5. No Quadro 5, visualiza-se o resultado da neutralização, no nível fonológico da representação: cancela-se o contraste entre as vogais médias altas e baixas pela desassociação dos traços do *tier* 5.

Quadro 5 – Níveis de representação, marcação e variantes de /e o/ em sílaba pretônica em PB.

Níveis de representação	/e/	/ɛ/	/o/	/ɔ/
<b>Fonológico</b> (Nível das Dimensões: oposições contrastivas)	1.[vogal] 2.[soante] 3.Impul.da.L. 4.Dorsal <sup>AL</sup> 5.Altura.da.L.	1.[vogal] 2.[soante] 3.Impul.da.L. 4.Dorsal <sup>AL</sup> 5.Dorsal <sup>AL</sup>	1.[vogal] 2.[soante] 3.Dorsal <sup>HL</sup> 4.Dorsal <sup>AL</sup> 5.Altura.da.L.	1.[vogal] 2.[soante] 3.Dorsal <sup>HL</sup> 4.Dorsal <sup>AL</sup> 5.Dorsal <sup>AL</sup>
	<b>Neutralização</b> 1.[vogal] 2.[soante] 3.Impul.da.L. 4.Dorsal <sup>AL</sup>	<b>Neutralização</b> 1.[vogal] 2.[soante] 3.Impul.da.L. 4.Dorsal <sup>AL</sup>	<b>Neutralização</b> 1.[vogal] 2.[soante] 3.Dorsal <sup>HL</sup> 4.Dorsal <sup>AL</sup>	<b>Neutralização</b> 1.[vogal] 2.[soante] 3.Dorsal <sup>HL</sup> 4.Dorsal <sup>AL</sup>
<b>Fonético-fonológico</b> (Nível dos Gestos: completações e aperfeiçoamentos)	<b>Alçamento:</b> OU completação com gesto [alto]		<b>Abaixamento:</b> OU completação com gesto [baixo]	
			Ø : inervação muscular neutra	
<b>Fonético</b> (Implementação: <i>continuum</i> de realizações)	i      u	ɛ      ɔ	e      o	

Fonte: Elaborado pela autora.

O Quadro 5 permite visualizar, além disso, que, neutralizado o contraste entre as vogais médias altas e baixas no nível fonológico, a completação de 4.Dorsal<sup>AL</sup> no nível fonético-fonológico pode ser com gesto [alto], quando a vogal é implementada como [i, u] no nível fonético (alçamento); ou com gesto [baixo], quando a vogal é implementada como [ɛ, ɔ] no nível fonético (abaixamento). Não havendo a completação de 4.Dorsal<sup>AL</sup> com um dos dois gestos no par antagônico, a inervação muscular é neutra e a vogal realiza-se como [e, o].

Purnell, Raimy e Salmons (2019) não propõem um modelo de representação para a completação e o aperfeiçoamento no nível fonético-fonológico. Usam o formato das regras de reescrita de Chomsky e Halle (1968), que eles designam ‘semiformais’, para assinalar aspectos relevantes das associações em questão. Segundo Purnell, Raimy e Salmons (2019, p. 461), sobre um processo como o *umlaut* (assimilação, pela vogal da raiz, do timbre da vogal do sufixo flexional) de [i] no inglês antigo: “as regras de reescrita do SPE [*Sound Pattern of English*, de Chomsky e Halle, 1968] fornecem um modo conveniente de chamar atenção para aspectos importantes da representação deste processo.” Nesse sentido, lança-se mão desse formato para destacar pontos de interesse nos processos aqui enfocados tanto na neutralização, que ocorre no nível fonológico, quanto nas completações fonético-fonológicas, que resultam em alçamento ou abaixamento das vogais médias pretônicas.

Na Figura 8, destaca-se que a neutralização do contraste entre as médias corresponde ao cancelamento da marcação Altura da Língua de último *tier* (*tier* 5). A neutralização afeta vogais marcadas Dorsal<sup>AL</sup> no *tier* 4, em sílaba que precede a sílaba acentuada<sup>9</sup>.

<sup>9</sup> ‘acento’ aparece entre aspas simples porque se reconhece tratar-se de uma propriedade atribuída pela estrutura métrica, não pela marcação com traços.

$$[5. \text{ Altura da língua}] \rightarrow \emptyset / \begin{array}{c} \text{---} \\ | \\ \left( \begin{array}{c} 1. \text{ vogal} \\ \vdots \\ 4. \text{ Dorsal}^{\text{AL}} \end{array} \right) \end{array} \text{C}_0 \left( \begin{array}{c} 1. \text{vogal} \\ \text{'acento'} \end{array} \right)$$

Fonte: Elaborada pela autora.

Figura 8 – Neutralização do contraste entre as vogais médias /e ε/ e /o ɔ/ em sílabas pretônicas.

A Figura 9 destaca a completção da marcação nula superordenada Dorsal<sup>AL</sup> no *tier* 4 com gesto [alto], envolvida na implementação das vogais como [i, u] no nível fonético-fonológico.

$$\emptyset \rightarrow [\text{alto}] / \begin{array}{c} \text{---} \\ | \\ \left( \begin{array}{c} 1. \text{vogal} \\ \vdots \\ 4. \text{Dorsal}^{\text{AL}} \end{array} \right) \end{array} \text{C}_0 \left( \begin{array}{c} 1. \text{vogal} \\ \text{'acento'} \end{array} \right)$$

Fonte: Elaborada pela autora.

Figura 9 – Alçamento das vogais médias pretônicas.

A Figura 10 assinala que, se a marcação nula superordenada Dorsal<sup>AL</sup> no *tier* 4 das vogais médias pretônicas é completada com o gesto [baixo] no nível fonético-fonológico, a vogal implementada é [ε, ɔ].

$$\emptyset \rightarrow [\text{baixo}] / \begin{array}{c} \text{---} \\ | \\ \left( \begin{array}{c} 1. \text{vogal} \\ \vdots \\ 4. \text{Dorsal}^{\text{AL}} \end{array} \right) \end{array} \text{C}_0 \left( \begin{array}{c} 1. \text{vogal} \\ \text{'acento'} \end{array} \right)$$

Fonte: Elaborada pela autora.

Figura 10 – Abaixamento das médias pretônicas: especificação com marcação neutralizada.

O que consta nas figuras 8, 9 e 10 vale como uma descrição heurística da interação de forças mente-corpo na emergência de padrões fonológicos, não como um modelo de derivação fonológica orientado pelo input. Mas serve para assinalar, especificamente sobre o alçamento, que a completção de 4. Dorsal<sup>AL</sup> com gesto [alto] é um de dois processos passíveis de se aplicar no nível fonético-fonológico a vogais neutralizadas em português.

No PB de Porto Alegre, onde as proporções de alçamento são baixas, menores para /e/ do que para /o/, a completção de 4. Dorsal<sup>AL</sup> com o gesto [alto] não é preponderante, mas possível. Como afirmamos em Battisti, Perozzo e Cunha (2020), quando se observa alçamento (completção com [alto]), ele resulta de múltiplos fatores: pode ser efeito coarticulatório, por compatibilização articulatória entre consoante coronal/dorsal e vogal; pode resultar também de redução por dispersão máxima de elementos no espaço vocálico, embora, como explicamos naquele trabalho, a redução vocálica possa ser refreada pelo “fato de /a/ pretônico, nessa variedade de PB, ser produzido em um ponto do espaço vocálico foneticamente mais próximo a [ε, ɔ], o que faz de [ε, ɔ] uma melhor opção de realização das médias pretônicas em termos de distâncias relativas” (Battisti; Perozzo; Cunha, 2020).

Ou seja, a completção, ou não, das dimensões contrastivas por um dos gestos de cada par resulta da complexa interação de forças que contribuem para “formatar a

aquisição, armazenamento de longo prazo e implementação *online* de conhecimento fonológico – forças emanando das áreas da percepção auditiva, produção da fala, memória lexical, morfossintaxe e assim por diante” (Harris, 2007, p. 137). No alçamento, além da coarticulação e da dispersão vocálica, pesa, por exemplo, a memória lexical. Brescancini et al. (2017) encontram, em dados de Porto Alegre do VARSUL, evidências de que o alçamento de /e/ vincula-se a certos itens lexicais (*senhora*) e radicais (*pequen-*), e o alçamento de /o/ é afetado pela frequência lexical, por força de itens lexicais pertencentes sobretudo à classe dos verbos (destacando-se os radicais *com-*, *começ-*, *conheç-* e *convers-*).

## 5 CONCLUSÃO

A análise empreendida neste artigo norteou-se por dois objetivos: buscar motivações representacionais para as proporções de alçamento observadas e representar os contrastes vocálicos do PB em um modelo de traços privativos pouco explorado em análises fonológicas do português, contribuindo para esclarecer suas potencialidades e limites.

No que se refere ao primeiro objetivo, a análise, coerente com estudos sobre a aquisição dos contrastes vocálicos do PB e dos processos variáveis que afetam as vogais médias em sílabas pretônicas (Matzenauer; Miranda, 2007; Matzenauer, 2009), estabelece a hierarquia contrastiva das vogais do PB, como fez antes Lee (2008, 2012), mas com o modelo de traços privativos de Purnell e Raimy (2015) e Purnell, Raimy e Salmons (2019). A análise mostra que /o/ é uma vogal menos marcada do que /e/, o que explica sua maior susceptibilidade ao alçamento no PB de Porto Alegre e, potencialmente, nas demais variedades de português.

Distinguindo, como propõe o modelo, o nível fonológico (das dimensões contrastivas) do nível fonético-fonológico (das completções por gestos) e do fonético (das implementações), a análise mostra que o alçamento variável se vincula tanto ao nível fonológico, por sua relação com a neutralização e com a marcação dos contrastes das médias, quanto à atividade no nível fonético-fonológico pela completção, ou não, da dimensão Dorsal<sup>AL</sup> no *tier* 4 por [alto]. Assim, o alçamento e outros processos de variação fonológica encontram, na abordagem fonológica de contraste e aperfeiçoamento modelada por Purnell e Raimy (2015) e Purnell, Raimy e Salmons (2019), um tratamento adequado em suas motivações estruturais e no que diz respeito aos processos a eles relacionados. Além disso, o uso, na análise, da estrutura implicacional de organização dos traços privativos proposta pelo modelo, tanto explica quanto limita a atividade fonológica e as modificações fônicas possíveis na língua: na neutralização, a dimensão Altura da Língua afetada é a do *tier* 5, o que exclui /i, u/ dos segmentos-alvo; e o alçamento pode afetar apenas segmentos Dorsal<sup>AL</sup> no *tier* 4.

Sobre o segundo objetivo, cabe acrescentar, às potencialidades afirmadas acima, que seguir o modelo de Purnell e Raimy (2015) e Purnell et al. (2019) não implica operar com propriedades contrastivas pré-determinadas, numa concepção de processamento orientada ao *input*. É um modelo que representa, com categorias abstratas, propriedades fônicas utilizadas para diferenciar significados nas línguas do mundo, internalizadas a partir da atividade fonológica. Assim fazendo, prevê a variabilidade de superfície, relacionando representação fonológica e processos fonético-fonológicos de forma compatível com teorias gerativistas (geometria de

traços) e argumentos como os exemplaristas (sobre a frequência dos itens lexicais e *continuum* de realizações no nível fonético).

A análise deixa uma questão em aberto, referente à repetição da dimensão Altura da Língua na marcação contrastiva das vogais médias. Temos dúvida sobre a adequação teórica dessa repetição, que se mostra necessária, na presente análise, para marcar o contraste de /ɛ, ɔ/ com /e, o/ na quarta (e última) iteração. É uma questão que deve merecer atenção em outros estudos para sustentar ou rever a repetição. Outra perspectiva de desenvolvimento deste artigo pode ser a realização de uma comparação sistemática da presente análise com outras referentes ao sistema vocálico do PB, concebidas conforme modelos distintos, como a de Mazzaferro e Matzenauer (2018), desenvolvida conforme Calabrese (2005), de marcação parametrizada dos contrastes vocálicos por uma Escala de Robustez de traços fonológicos. Essas são possibilidades de desdobramento do trabalho, que o artigo deixa como perspectivas futuras de investigação.

## REFERÊNCIAS

- Avery P, Idsardi W. Laryngeal dimensions, completion and enhancement. In: Hall TA, editor. *Distinctive feature theory*. Berlin: Mouton de Gruyter; 2001. p. 41-70.
- Battisti E, Cunha, VG. A representação dos contrastes vocálicos e a elevação não harmônica das vogais médias pretônicas em uma variedade de português brasileiro. XIII Congresso Internacional da ABRALIN; 30 out-03 nov. 2023; Curitiba, PR, Brasil.
- Battisti E, Perozzo RV, Cunha VG. Alçamento sem motivação aparente em uma variedade de português brasileiro: efeitos de coarticulação de sequências CV e dispersão vocálica. *Entrepalavras*. 2020;10(1):12-35. [citado 13 jun. 2024]. Disponível em: <https://doi.org/10.22168/2237-6321-11757>.
- Battisti E, Soares EPM. Representação fonológica e variabilidade de superfície da lateral palatal no português brasileiro. *Letrônica*. 2022;15(1):1-17. [citado 13 jun. 2024]. Disponível em: <https://doi.org/10.15448/1984-4301.2022.1.42522>.
- Bermúdez-Otero R. Phonological change in Optimality Theory. In: Brown K, editor. *Encyclopedia of language and linguistics*. 2<sup>nd</sup> ed. Oxford: Elsevier; 2006. p. 497–505.
- Bermúdez-Otero R. Amphichronic explanation and the life cycle of phonological processes. In: Honeybone P, Joseph Salmons J, editors. *The Oxford handbook of historical phonology*. Online edition: Oxford Academic; 2013. [citado 13 jun. 2024]. Disponível em: <https://doi.org/10.1093/oxfordhb/9780199232819.013.014>.
- Biasibetti APCS. A aplicação variável da elevação sem motivação aparente das vogais médias pretônicas em Porto Alegre: considerações à luz da Teoria de Exemplares. *Letrônica*. 2014;7(2):496-521.
- Bisol L. O alçamento da pretônica sem motivação aparente. In: Bisol L, Collischonn G, organizadoras. *Português do Sul do Brasil: variação fonológica*. Porto Alegre: EDIPUCRS; 2009. p. 73-92.
- Brescancini CR, et al. Alçamento da vogal pré-tônica em Porto Alegre-RS: léxico e variação. *ReVEL*. 2017;14(Nº esp):4-24.
- Calabrese A. *Markedness and economy in a derivational model of phonology*. Mouton de Gruyter: Berlin; 2005.
- Camara Jr. JM. *Estrutura da Língua Portuguesa*. Petrópolis: Editora Vozes; 1970.
- Chomsky N, Halle M. *The sound pattern of English*. New York: Harper and Row; 1968.
- Clements GN. Vowel height assimilation in Bantu languages. *Working papers of the Cornell Phonetics Laboratory*. 1991;5:37-76.

- Coetzee AW. A comprehensive model of phonological variation: grammatical and non-grammatical factors in variable nasal place assimilation. *Phonology*. 2016;33(2):211-146. [citado 13 jun. 2024]. Disponível em: doi:10.1017/S0952675716000117.
- Cruz MC. As vogais médias pretônicas em Porto Alegre/RS: um estudo sobre o açamento sem motivação aparente [dissertação]. Porto Alegre: Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul; 2010.
- Cunha VG. Estudo-piloto da percepção e avaliação da elevação sem motivação aparente por porto-alegrenses. XXX Salão de Iniciação Científica da UFRGS; 18-22 out. 2018; Porto Alegre, RS, Brasil.
- Cunha VG. A elevação variável das vogais médias pretônicas não desencadeada por vogal alta seguinte no português de Porto Alegre: da produção à percepção e avaliação linguística [trabalho de conclusão de curso]. Porto Alegre: Instituto de Letras, Universidade Federal do Rio Grande do Sul; 2022.
- Dresher E. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press; 2009.
- Goldstein L, Byrd D, Saltzman E. The role of vocal tract gestural action units in understanding the evolution of phonology. In: Arbib M, editor. *Action to language via the mirror neuron system*. Cambridge: Cambridge University Press; 2006. p. 215-249.
- Hall D. Phonological contrast and its phonetic enhancement: dispersedness without dispersion. *Phonology*. 2011;28(1):1-54.
- Harris J. Representation. In: de Lacy P, editor. *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press; 2007. p. 119-137.
- Kiparsky P. Labov, sound change and phonological theory. *Journal of Sociolinguistics*. 2016;20(4):464-488.
- Klunck P. Açamento das vogais médias pretônicas sem motivação aparente [dissertação]. Porto Alegre: Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul; 2007.
- Labov W. *Sociolinguistic patterns*. Philadelphia: University of Philadelphia Press; 1972.
- Labov W. Resolving the neogrammarian controversy. *Language*. 1981;57(2):267-308.
- Labov W. *Principles of linguistic change*. Oxford: Blackwell; 1994. (Vol. I: Internal factors).
- Lee S-H. Contraste das vogais no PB. *Portuguese-Brazilian studies*. 2008;5:201-221.
- Lee S-H. Variação fonológica e contraste no sistema vocálico do PB. In: Lee S-H, organizador. *Vogais além de Belo Horizonte*. Belo Horizonte: FALE/UFGM; 2012. p. 7-16.
- Lee S-H, Oliveira MA. Variação inter- e intradialetal no PB: um problema para a teoria fonológica. In: Hora D da, Collischonn G, organizadores. *Teoria linguística: fonologia e outros temas*. João Pessoa: Editora da UFPB; 2003. p. 67-91.
- Lindblom B. Phonetic universals in vowel systems. In: Ohala J, Jaeger J, editores. *Experimental phonology*. Orlando: Academic Press; 1986. p. 13-44.
- Maddieson I. *Patterns of sounds*. Cambridge: Cambridge University Press; 1984.
- Matzenauer CLB, Miranda ARM. Traços distintivos e aquisição das vogais do PB. *SIS Vogais*; 15-17 nov. 2007; João Pessoa, PB, Brasil.
- Matzenauer CLB. Sobre as vogais médias pretônicas na aquisição do português brasileiro. *Organon*. 2009;29(46):71-108.
- Mazaferro GT, Matzenauer CLB. Oposições entre segmentos: vogais médias na aquisição e em tipologias de línguas. *Linguagem & Ensino*. 2018;21(1):129-156.
- McCarthy J, Prince A. Faithfulness and reduplicative identity. In: Beckman J, et al., editores. *Papers in Optimality Theory – UMass Occasional Papers 18*. Amherst, Massachusetts: GLSA; 1995. p. 249-384.
- Monaretto VN. de O. O açamento das vogais médias pretônicas /e/ e /o/ sem motivação aparente: um estudo em tempo real. *Fragmentum*. 2013;39:19-29.

- Natvig D. Rhotic underspecification: deriving variability and arbitrariness through phonological representations. *Glossa: a journal of general linguistics*. 2020;5(1):1–28. [citado 13 jun. 2024]. Disponível em: <https://doi.org/10.5334/gjgl.1172>.
- Prince A, Smolensky P. *Optimality Theory: constraint interaction in generative grammar*. Rutgers University/University of Colorado: Boulder; 1993. Relatório técnico.
- Purnell TC, Raimy E. Distinctive features, levels of representation, and historical phonology. In: Honeybone P, Salmons J, editores. *The Handbook of Historical Phonology*. Oxford: Oxford University Press; 2015. p. 522–544.
- Purnell TC, Raimy E, Salmons J. Old English vowels: diachrony, privativity, and phonological representations. *Language*. 2019;95(4):e447-e473. [citado 13 jun. 2024]. Disponível em: <https://muse.jhu.edu/article/743116/pdf>.
- R Core Team. *R: a language and environment for Statistical Computing* [programa de computador]. Viena: R Foundation for Statistical Computing; 2021. [citado 13 jun. 2024]. Disponível em: <https://www.r-project.org/>.
- Silva APC. *Elevação sem motivação aparente das vogais médias pretônicas entre os jovens porto-alegrenses* [dissertação]. Porto Alegre: Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul; 2014.
- Silva MB. *Vogais pretônicas no Brasil: uma proposta de descrição a partir da fala de Salvador*. São Paulo: Blucher; 2021.
- Spahr C. *Contrastive representations in non-segmental phonology* [dissertação]. Toronto: University of Toronto; 2016.
- Wetzels WL. The representation of vowel height and vowel height neutralization in Brazilian Portuguese (Southern Dialects). In: Goldsmith JA, Hume E, Wetzels WLM, editores. *Tones and features: phonetic and phonological perspectives*. Berlin: Walter de Gruyter; 2011. p. 331-359.



# Os efeitos do aprendizado das letras <e, o> na produção oral de vogais médias pretônicas e postônicas finais

## *The effects of learning the letters <e, o> on the oral production of mid vowels in pre-stressed and final post-stressed positions*

Cecília Toledo\*

*Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil*

**Resumo:** Este artigo investiga a influência do aprendizado das letras <e, o> na produção variável de vogais médias e altas pretônicas (ex: [me'ninɔ] ~ [mi'ninɔ]) e postônicas finais (ex: [pato] ~ [patɔ]) na fala do português brasileiro. Estudos já documentaram que, ao longo da escolarização, a escrita pode retroalimentar a fala (Chevrot; Beaud; Varga, 2000; Schiwindt et al., 2007; Adamoli, 2012). O presente texto avança em relação à literatura precedente por observar os efeitos da escrita na fala em dois contextos acentuais distintos: pretônico e postônico final. Foram coletados, experimentalmente, dados de fala e de escrita de alunos do 1º, 3º, 5º, 7º e 9º anos do Ensino Fundamental de Belo Horizonte - MG. Os resultados mostraram que os aprendizes aumentam a pronúncia de vogais médias, em relação à pronúncia de vogais altas, em contexto pretônico, mas não em contexto postônico final. Ancorando-se nos Modelos de Exemplares (Johnson, 1997; Bybee, 2001; Pierrehumbert, 2001), argumentamos que o conhecimento ortográfico pode auxiliar, probabilisticamente, na organização do conhecimento linguístico, contribuindo para o fortalecimento de categorias fonológicas. Contudo, a ortografia opera de maneira específica em cada categoria, tendo em vista que as categorias fonológicas são múltiplas, probabilísticas e dinâmicas.

**Palavras-chave:** Oralidade. Escrita. Vogais médias. Representação probabilística.

**Abstract:** This paper investigates the influence of learning the letters <e, o> on the variable production of mid vowels in spoken Brazilian Portuguese, focusing on pre-stressed (e.g.: [me'ninɔ] ~ [mi'ninɔ]) and final post-stressed syllables (e.g.: [pato] ~ [patɔ]). Numerous studies have reported the reciprocal influence between oral production and writing across school years (Chevrot; Beaud; Varga, 2000; Schiwindt et al., 2007; Adamoli, 2012). Building upon prior research on the topic, this study advances by examining the interplay between spoken and written language in two different stress positions: pre-stressed and final post-stressed syllables. An experiment was carried out to collect spoken and written data from students in the 1st, 3rd, 5th, 7th, and 9th school years in Belo Horizonte, Brazil. Results showed increased production rates of mid vowels in comparison to high vowels in pre-stressed position, but not in in final post-stressed position. Drawing on Exemplar Models (Johnson, 1997; Bybee, 2001; Pierrehumbert, 2001), we posit that orthographic knowledge may probabilistically contribute to organizing the linguistic knowledge, by strengthening phonological categories. However, orthography's impact is expected to vary uniquely in each category, since phonological categories are multifaceted, probabilistic, and dynamic.

**Keywords:** Speech. Spelling. Mid vowels. Probabilistic representation.

---

\* Professora substituta de Língua Portuguesa, Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil; [ceciliavstoledo@gmail.com](mailto:ceciliavstoledo@gmail.com)

## 1 INTRODUÇÃO

Um fato recorrentemente documentado no português brasileiro é o de que aprendizes de escrita podem cometer erros ortográficos envolvendo vogais médias e altas pretônicas e postônicas finais (Miranda, 2006). Por exemplo: <\*pirigu> (*perigo*), <\*boneca> (*boneca*), <\*penti> (*pente*), <\*patu> (*pato*). Esse tipo de erro ortográfico é, comumente, explicado como reflexo da oralidade. Isto é, fenômenos fonológicos, que motivam pronúncias como [pi'riɡũ] e [pẽ̃ʎi], alimentam a escrita, promovendo trocas entre as letras <e ~ i> e <o ~ u>. Os reflexos da oralidade na escrita já foram amplamente investigados em trabalhos como os de Alvarenga et al. (1989), Lemos (2001), Miranda (2006, 2008, 2010) e Tenani e Reis (2011), para citar alguns.

Achados mais recentes têm ampliado os estudos sobre oralidade e escrita mostrando que não só a fala alimenta a escrita, mas também a escrita pode retroalimentar a fala (Chevrot; Beaud; Varga, 2000; Schwindt et al., 2007; Cristófarosilva; Greco, 2010; Adamoli, 2012). Isto é, o aprendizado da forma ortográfica pode motivar alterações nas pronúncias de determinadas palavras. Sabendo-se disso, este artigo se propõe a investigar a influência do aprendizado das letras <e, o> na variação oral de vogais médias e altas pretônicas e postônicas finais em português brasileiro. Em outros termos, busca-se avaliar se o aprendizado da ortografia de palavras como <mEnino> e <pentE>, por exemplo, motiva o aumento de pronúncias como [me'nino] e [pẽ̃tẽ] em detrimento de [mi'ninũ] e [pẽ̃ũ].

O presente trabalho avança em relação à literatura precedente por observar os efeitos da forma escrita na forma falada em dois contextos acentuais distintos: o contexto pretônico e o contexto postônico final. A seguinte hipótese de pesquisa foi testada: *o aprendizado das letras <e, o> motiva o aumento de vogais médias na fala a depender do contexto acentual: pretônico ou postônico final.*

Essa hipótese é fundamentada, sobretudo, nas seguintes premissas dos Modelos de Exemplares (Johnson, 1997; Bybee, 2001, 2002; Pierrehumbert, 2001, 2003; Foulkes; Docherty, 2006; Port, 2007):

- a) as unidades linguísticas armazenadas na memória humana são unidades atestadas no uso real da língua;
- b) as categorias linguísticas são múltiplas, pois armazenam informações fonéticas, lexicais, contextuais, semânticas, socioculturais etc.;
- c) as categorias linguísticas se organizam de forma probabilística;
- d) o conhecimento linguístico é plástico e dinâmico.

Ancorando-se nas premissas a), b), c) e d), este artigo visa argumentar que o conhecimento gráfico, adquirido ao longo dos anos escolares, pode influenciar, probabilisticamente, na organização da representação fonológica. Isto é, na medida em que os indivíduos vivenciam experiências gráficas, eles vão construindo representações fono-ortográficas. Assim, um padrão ortográfico – como as letras <e, o> pretônicas e postônicas finais – pode se relacionar às categorias sonoras e contribuir para o fortalecimento de determinadas representações. Além disso, argumenta-se que a escrita pode interferir em maior ou menor grau na sonoridade da língua, a depender, sobretudo, do tipo de categoria fonológica avaliada. Especificamente, buscamos defender que os efeitos de um padrão ortográfico na fala e no sistema linguístico variam a depender do contexto acentual.

O artigo está organizado da seguinte maneira: na próxima seção, apresentamos alguns trabalhos que já registraram o efeito retroalimentador da escrita na fala. Na segunda seção, revisamos pesquisas que documentaram a variação sonora em contextos pretônico e postônico final. Na terceira, apresentamos alguns pressupostos teóricos que sustentam a discussão proposta neste texto. Na quarta, descrevemos a metodologia desta pesquisa. Na quinta, analisamos os dados e discutimos teoricamente os resultados. Na sexta e última seção, apresentamos as considerações finais.

## 2 OS EFEITOS DA ESCRITA NA FALA

A reflexão sobre a língua materna é uma das habilidades cruciais para o desenvolvimento da escrita alfabética. Isso porque os alfabetos são, essencialmente, sistemas formados por um conjunto finito de símbolos gráficos que representam os símbolos fonológicos (Man, 2002; Veloso, 2005; Soares, 2018); ou seja, para compreender o princípio alfabético é necessário se tornar capaz de relacionar letras (grafemas) e sons (fonemas). Como afirma Oslon (1999, p. 32), “aprender a escrever é aprender a ouvir e a pensar sobre a sua própria língua de uma nova maneira.” O aprendiz de escrita precisa ser capaz de acessar as representações sobre sua língua materna (Miranda, 2006) e, mais do que isso, precisa fazer com que o conhecimento tácito sobre a linguagem passe a ser explícito (Gombert, 2003; Seymour, 2005). De forma geral, o sucesso no aprendizado da escrita depende do acesso à gramática e do desenvolvimento da consciência metalinguística, que se caracteriza como “a habilidade de refletir sobre e manipular os aspectos estruturais da língua” (Nagy; Anderson, 1999, p. 155).

Sabendo-se, então, da íntima relação entre os conhecimentos oral e gráfico, pode-se dizer que a análise da escrita inicial pode trazer ricas evidências sobre a reflexão dos aprendizes quanto à língua materna. Como bem define Miranda (2008, p. 152), “a escrita inicial das crianças é uma fonte extraordinária de dados, que, ao serem analisados e interpretados, podem contribuir para uma melhor compreensão a respeito dos conhecimentos que elas já possuem sobre a língua materna.” É sabido, e muito bem documentado na literatura da área (Lemle, 1987; Alvarenga et al., 1989; Lemos, 2001; Miranda, 2010; Nóbrega, 2013), que a escrita infantil reflete bastante oralidade.

Para além dos efeitos da oralidade na escrita, há algumas evidências de que o aprendizado da escrita alfabética pode modificar as formas de uso da língua falada (Chevrot; Beaud; Varga, 2000; Schiwindt et al., 2007; Cristófaros-Silva; Greco, 2010; Adamoli, 2012). Kato (1990) distingue dois tipos de fala: a que antecede o letramento e a que o sucede; ou seja, a fala pré-letrada e a fala pós-letrada. Para o autor, a língua falada por indivíduos altamente alfabetizados em sociedades letradas é fortemente influenciada pela língua escrita. Por outro lado, a fala de indivíduos pré-alfabetizados é muito pouco afetada pelo sistema gráfico (Kato, 1990, p. 11). Em busca de discutir em maiores detalhes os efeitos do aprendizado da escrita na fala, esta seção se concentrará, daqui em diante, à revisão de alguns dos trabalhos sobre esse tema.

Em uma pesquisa sociolinguística sobre a aquisição de variáveis sonoras, Chevrot, Beaud e Varga (2000) formularam a seguinte hipótese: o contato com palavras escritas é um fator que influencia na aquisição de unidades fonológicas variáveis. Para testar a hipótese, os autores analisaram a pronúncia do R pós-consonantal do francês – como em *sucrer* ‘açúcar’ – na fala de crianças em fase de alfabetização. O R pós-consonantal é altamente cancelado na fala de adultos franceses.

Assim, uma palavra como *sucrer* é, recorrentemente, pronunciada como [syk] – suc(re) por adultos. Os autores testaram se, após terem o contato com o <r> ortográfico, as crianças francesas aumentariam os índices de produção da variante [r], ou se o cancelamento prevaleceria na fala infantil. Os resultados indicaram que o aprendizado da ortografia motivou o aumento da pronúncia do rótico na fala das crianças. Para os autores, esses resultados indicam que “a representação ortográfica de uma palavra permite que as crianças complementem a representação fonológica com um fonema que é frequentemente excluído no ambiente linguístico.” (Chevrot; Beaud; Varga, 2000, p. 321, tradução nossa). Na visão de Chevrot, Beaud e Varga (2000), a forma escrita atualizaria a representação fonológica dos indivíduos.

Adamoli (2012) também propôs que a forma escrita pode atualizar a gramática, após verificar a influência da escrita na pronúncia de ditongos variáveis. O autor analisou a variação dos ditongos [aɪ] e [eɪ] na fala infantil. Exemplo: [ˈfaɪʃə] ~ [ˈfaʃə] ‘faixa’ e [ˈkeɪʒu] ~ [ˈkeʒu] ‘queijo’. As crianças concluíram o 1º ano escolar falando e escrevendo formas monotongadas, como \**quejo* e \**faxa*, em maiores índices. Ao final do 2º ano, as crianças aumentaram os índices de produção oral e escrita dos ditongos – como em *queijo* e *faixa*. O autor concluiu que a estrutura do ditongo (vogal + glide) é adquirida pelas crianças por meio do aprendizado da forma ortográfica.

Chevrot, Beaud e Varga (2000) e Adamoli (2012) explicam as influências da escrita na fala com o suporte de modelos formais da fonologia. Diferentemente, a discussão feita neste artigo é ancorada nas premissas dos Modelos de Exemplares (Johnson, 1997, 2005; Bybee, 2001, 2002, 2010; Pierrehumbert, 2001, 2003; Foulkes; Docherty, 2006; Port, 2007). Ancorando-se na ideia de que o sistema linguístico é dinâmico e probabilístico, propõem-se que a escrita pode contribuir com o aumento da robustez de determinadas categorias sonoras, o que pode motivar o aumento de determinadas variantes orais. Esta proposta dialoga com a apresentada por Cristófaros-Silva e Greco (2010). As autoras reportaram resultados da pesquisa de Greco (2009) sobre as influências da forma ortográfica na pronúncia de vogais pretônicas. Os resultados reportados mostraram que as crianças do 3º (62%) e 5º (52%) anos do EFI pronunciaram a vogal [e] – [meˈninu] – em maiores índices do que as crianças do 1º ano (30%); isto é, os alunos dos anos escolares mais avançados tenderam a pronunciar as vogais da forma mais próxima à representada pela escrita. A partir desses resultados, Cristófaros-Silva e Greco (2010, p. 91) sugeriram que as representações fonológicas podem ser reorganizadas ao longo do tempo, devido às pressões sociais, como a escolarização. O presente trabalho também reflete sobre a natureza da representação mental a partir da análise da influência da escrita na fala, mas avança ao analisá-la em dois contextos linguísticos: vogais pretônicas e vogais postônicas finais. Como veremos a seguir, as vogais pretônicas e postônicas finais operam de forma específica no sistema fonético-fonológico da língua portuguesa, o que justifica a hipótese de que o aprendizado da escrita motiva diferentes efeitos na produção oral dessas vogais.

### 3 A VARIAÇÃO DAS VOGAIS PRETÔNICAS E POSTÔNICAS FINAIS NO PORTUGUÊS BRASILEIRO

A descrição oral de segmentos vocálicos é feita de acordo com suas posições acentuais. As vogais podem ser tônicas, pretônicas e postônicas. As vogais tônicas

ocorrem em sílaba portadoras do acento primário<sup>1</sup>: *pecado*. As vogais pretônicas ocorrem em sílabas que antecedem a tônica – *pecado* – e as postônicas, em sílabas que sucedem a tônica – *pecado* (Cristófaros-Silva, 2013). Neste artigo, ênfase é dada nas vogais pretônicas e postônicas finais.

Uma diferença explícita entre as vogais pretônicas e postônicas finais é o contexto/ambiente em que ocorrem. As vogais pretônicas podem ocorrer em início e/ou meio da palavra, em sílabas próximas ou distantes da tônica, em mais de uma sílaba da mesma palavra, próximas de vogais e/ou consoantes (Cristófaros-Silva; Guimarães, 2013). Por exemplo: *leão*, *inseguro*, *buraco*, *paralelepípedo*. Por outro lado, as vogais postônicas finais ocorrem, necessariamente, em posição final de palavras. Exemplo: *pato*, *dente*, *gata*. Ou seja, no que diz respeito ao contexto fonético, as vogais pretônicas são mais heterogêneas do que as vogais postônicas finais (Cristófaros-Silva; Guimarães, 2013).

O caráter heterogêneo das vogais pretônicas pode ser evidenciado também na distribuição fonética. Na variedade de Belo Horizonte – analisada neste artigo – algumas palavras são pronunciadas apenas com vogais altas pretônicas – *b[i]scoito*, *b[u]raco*; outras ocorrem regularmente com vogais médias pretônicas – *m[e]tade*, *p[o]der*. Há também palavras envolvidas em variação sonora, que são ora pronunciadas com vogais altas, ora com médias – *m[e]nino* ~ *m[i]nino*; *f[o]rmiga* ~ *f[u]rmiga* (Oliveira, 1992).

Tradicionalmente, a variação no contexto pretônico é explicada como resultado do fenômeno de *alçamento vocálico*, que ocorre quando vogais médias passam a ser pronunciadas como vogais altas (Viegas, 1995). Muitos trabalhos se dedicaram a descrever tal fenômeno no PB, o que possibilitou uma rica e minuciosa descrição dessas vogais no contexto referido (Bisol, 1981; Schiwindt, 1995; Viegas, 1995; Chaves, 2013; Magalhães; Campos, 2022).

Bisol (1981) propôs que o alçamento das vogais pretônicas é motivado por um outro fenômeno, denominado de harmonia vocálica, que ocorre quando a pronúncia de uma vogal é alterada por influência de outra vogal da mesma palavra (Crystal, 1988). Bisol (1981) percebeu que, no PB, as vogais médias pretônicas se realizam como altas, quando a vogal tônica é alta. Por exemplo: *m[e]nino* > *m[i]nino*. A autora propôs que o processo de harmonia vocálica no contexto pretônico do PB se realiza de maneira variável; isto é, mesmo em ambientes propícios, o alçamento da vogal pretônica ora ocorre, ora não.

Questionando a proposta de Bisol (1981), Oliveira (1992) demonstrou que a harmonia vocálica se aplica em apenas parte do léxico do PB e nem sempre há variação. No português de Belo Horizonte - MG, por exemplo, a palavra [mi.'ʃidə] *metida* é regularmente pronunciada com [i] pretônico, enquanto [me.'dʒi. tə] *medita* se realiza com [e] na primeira sílaba (Oliveira, 1992, p. 33). O autor levantou evidências de que a variação das vogais pretônicas tem motivação lexical. Isso significa que cada item lexical se comporta de um jeito diante do fenômeno de alçamento vocálico: uns são produzidos sistematicamente com vogais altas pretônicas; outros apresentam flutuações entre vogais médias e altas, e alguns se realizam com vogais médias. Para Oliveira (1992), apenas as palavras envolvidas em flutuação podem ser consideradas casos legítimos de variação sonora.

<sup>1</sup> Principal proeminência da palavra (Cristófaros-Silva, 2017).

Diferentemente das vogais pretônicas, as postônicas finais estão envolvidas em um outro tipo de fenômeno fonológico denominado, geralmente, de redução vocálica (Assis, 2017). A redução vocálica pode ser definida como um fenômeno de enfraquecimento das vogais, que, em estágio avançado, pode resultar no cancelamento vocálico. Embora haja evidências de que as vogais pretônicas também podem reduzir (Souza, 2012), o fenômeno se mostra mais recorrente em contexto postônico final (Dias; Seara, 2013). A redução vocálica é variável no contexto postônico final, o que significa que o fenômeno ora ocorre, ora não. Quando as vogais postônicas finais são pronunciadas, elas ocorrem sistematicamente como [ɪ, ə, ʊ]. Por exemplo: ['mɔɪɪ] *mole*; ['balə] *bala*, ['patʊ] *pato*, na grande maioria das variedades do PB (Bisol, 2003).

Cristófaros-Silva e Vieira (2015) argumentam que o cancelamento das vogais postônicas finais é um padrão inovador na língua portuguesa. Tomando as autoras como referência, pode-se dizer que as vogais médias postônicas finais estão em estágios mais avançados do que as vogais pretônicas no percurso evolutivo da língua portuguesa.

De modo geral, o objetivo desta seção foi o de demonstrar que as vogais pretônicas e postônicas finais da língua portuguesa estão envolvidas em diferentes fenômenos fonológicos. Em contexto pretônico, há variação entre vogais médias e altas. Exemplo: f[o]rmiga ~ f[u]rmiga. Essa variação é, tradicionalmente, explicada pelo fenômeno de alçamento vocálico. Já em contexto postônico final, ocorre o fenômeno de redução vocálica, que pode ser resumido pela seguinte trajetória evolutiva: vogal média > vogal alta > cancelamento vocálico.

Embora haja diferenças na produção oral das vogais pretônicas e postônicas finais, há uma semelhança no que diz respeito à escrita dessas vogais: em ambos os contextos, se pode escrever as letras <e> e <o>. Vejamos alguns exemplos a seguir:

Quadro 1 – Exemplos de palavras com vogais pretônicas.

(a) <menino, tesoura, teatro>	(b) <formiga, coruja, mochila>
(a') <pente, alface, peixe>	(b') <pato, garfo, cavalo>

Fonte: Elaborado pela autora.

Em (a) e (a'), encontram-se, respectivamente, palavras escritas com a letra <e> em posição pretônica e postônica final. Em b e b', tem-se palavras grafadas com <o> pretônico e postônico final. Considerando-se as influências da escrita na fala, formulamos a seguinte questão: *o aprendizado das letras <e, o> motiva o aumento de vogais médias pretônicas e postônicas finais na oralidade?* Partindo do pressuposto de que as vogais pretônicas e postônicas finais operam de forma distinta no sistema linguístico do português brasileiro, formulamos a hipótese de que *o aprendizado das letras <e, o> motiva o aumento de vogais médias na fala a depender do contexto acentual: pretônico ou postônico final.* Vejamos, a seguir, as premissas teóricas que fundamentam essa hipótese.

#### 4 ESCOPO TEÓRICO

A discussão proposta neste artigo se ancora em alguns pressupostos dos Modelos de Exemplares (Johnson, 1997; Bybee, 2001, 2002; Pierrehumbert, 2001, 2003; Foulkes; Docherty, 2006; Port, 2007). Os Modelos de Exemplares partem da premissa de que as unidades linguísticas armazenadas na memória humana são

unidades atestadas no uso real da língua; isto é, os Modelos de Exemplares propõem que o conhecimento linguístico representa a totalidade de experiências que um indivíduo vivencia (Foulkes; Docherty, 2006). Nesta abordagem, fonemas e regras fonológicas são substituídos por categorias fonéticas altamente detalhadas. Cada categoria contém um grande conjunto de exemplares dos variados itens experienciados no uso linguístico. As representações são, portanto, altamente redundantes, pois uma mesma palavra pode estar registrada na memória muitas vezes, haja vista o grande número de vezes que uma palavra se repete no uso linguístico.

Vale ressaltar que há variações fonéticas na produção de itens lexicais com estruturas internas semelhantes. Oliveira (1992) mostrou que palavras com sílabas semelhantes, como *metida* e *medita* se comportam de forma distinta diante do fenômeno de alçamento vocálico no português de Belo Horizonte – MG. Bybee (2001), ao estudar o inglês americano, mostrou que palavras como *every*, *memory* e *mammary* variam de forma específica no que diz respeito à redução do *schwa*. Bybee (2002) mostrou que a frequência com que cada palavra ocorre na língua interfere no fenômeno de apagamento das consoantes [t] e [d] finais no inglês. Huback (2003) encontrou diferenças no que diz respeito ao cancelamento do R-final em nomes terminados pelo sufixo -dor. Fontes-Martins (2007) mostrou que uma mesma palavra pode ser produzida de forma distinta por diferentes indivíduos. Para os Modelos de Exemplares, todas essas variações encontradas na produção lexical são parte da representação mental.

Além da variação fonética, as categorias linguísticas propostas pelos Modelos de Exemplares armazenam informações contextuais, semânticas, socioculturais etc. (Johnson, 1997; Foulkes; Docherty, 2006). Em sociedades grafocênicas, a escrita pode ser o objeto sociocultural mais respeitado e mais utilizado no dia a dia. Livros são parte de nossa vida desde muito cedo. Antes mesmo de aprendermos a ler e a escrever, folheamos páginas cheias de letras (Soares, 2010). Até indivíduos analfabetos estão em contato com a escrita, tendo em vista que esta é parte da paisagem urbana de sociedades letradas. Nas ruas, presenciamos, com naturalidade, as placas, os outdoors, os letreiros, os panfletos etc. Sabendo-se que a escrita está tão presente em nossas vidas e, partindo da premissa de que as categorias linguísticas armazenam experiências socioculturais, é possível propor que a escrita influencia na construção do nosso conhecimento linguístico.

Para Port (2007), nossas intuições sobre a oralidade são fortemente influenciadas pelas formas ortográficas que aprendemos ao longo do tempo. Para Morais (1996), a alfabetização estabelece uma nova relação entre língua e falante de modo que não é mais possível ouvir uma palavra sem relacioná-la a sua forma ortográfica. Paiva (2003) concluiu que o comportamento do indivíduo muda na medida em que há o aumento do contato com a escrita. Para Mattingly (1992), a consciência linguística do falante é moldada pelo sistema ortográfico. Todas essas conclusões documentadas na literatura nos motivam a pensar que a forma escrita é parte das representações mentais. A premissa de multiplicidade da gramática, defendida pelos Modelos de Exemplares, tem potencial para englobar todos os subcampos da linguística – inclusive a escrita – na teoria da linguagem.

Vale ressaltar, ainda, que, para os Modelos de Exemplares, as categorias linguísticas se organizam de forma probabilística (Pierrehumbert, 2003). Isso significa que o número de vezes que um indivíduo experiencia determinada forma linguística

desempenha um papel importante na construção do conhecimento abstrato da língua. Além disso, o número de vezes que uma mesma forma linguística é produzida e percebida pelo falante pode afetar a probabilidade desta mesma palavra voltar a ocorrer no uso linguístico (Bod; Hay; Jannedy, 2003). Palavras frequentes no uso linguístico possuem muitos exemplares e, conseqüentemente, possuem representações robustas. Por outro lado, palavras pouco frequentes, com poucos exemplares armazenados na gramática, podem desaparecer, assim como outras memórias. A frequência com que uma palavra (frequência de ocorrência) ou um padrão (frequência de tipo) ocorre na experiência afeta a sua produtividade na língua (Bybee, 2001).

O termo produtividade “se refere à tendência de generalização de determinado padrão como reflexo da experiência linguística.” (Cristófaros-Silva; Guimarães, 2013, p. 318). Quanto mais itens são associados a uma mesma categoria, maior é sua produtividade; isto é, maior a possibilidade de um padrão geral se expandir para novos itens.

Além de ser probabilístico, o conhecimento linguístico é, nos Modelos de Exemplares, plástico e dinâmico (Cristófaros-Silva; Gomes, 2004). Esse pressuposto tem potencial para explicar as diferenças no comportamento linguístico de pessoas com diferentes graus de escolaridade. Morais et al. (1979), por exemplo, mostraram que indivíduos alfabetizados julgam a linguagem de forma diferente de indivíduos analfabetos. Esse fato traz evidências de que a gramática de indivíduos escolarizados armazena experiências distintas daquelas que não frequentaram a escola.

Sabendo-se que o comportamento de um mesmo indivíduo varia ao longo do tempo, pode-se pensar que os padrões linguísticos gerais, compartilhados pela comunidade de fala, emergem das experiências individuais. Nesse sentido, indivíduos alfabetizados podem agregar formas ortográficas às representações mentais, ampliando o grau de generalização das categorias. Por outro lado, os indivíduos pré-alfabetizados ou analfabetos constroem suas generalizações linguísticas por vias diferentes dos alfabetizados. Assim sendo, indivíduos com diferentes graus de escolarização compartilharão representações mentais diferentes, haja vista que vivem realidades diferentes. Em suma, os sistemas linguísticos se alteram ao longo do tempo, uma vez que são abertos às múltiplas experiências individuais. Nesse sentido, o aprendizado da ortografia, ao longo da escolarização, pode influenciar na organização do sistema linguístico, o que explica as interferências da escrita na fala.

Neste artigo, buscamos avaliar os efeitos do aprendizado das letras <e, o> na produção oral de crianças e de adolescentes belorizontinos. Argumentamos que o conhecimento ortográfico pode auxiliar, probabilisticamente, na organização do conhecimento linguístico, contribuindo para o fortalecimento de determinadas representações fonológicas e, conseqüentemente, para o aumento da produção oral de determinadas variantes orais. Isto é, na medida em que os indivíduos vivenciam experiências gráficas, eles vão construindo representações fono-ortográficas. Assim, um padrão ortográfico – como as letras <e, o> pretônicas e postônicas finais – pode se relacionar às variantes sonoras e contribuir para o fortalecimento das representações mentais. Como as representações são múltiplas, probabilísticas, plásticas e dinâmicas, é esperado que os padrões ortográficos exerçam diferentes influências no sistema linguístico, a depender de variados fatores, como o contexto acentual (pretônico ou postônico final), por exemplo.

## 5 METODOLOGIA

O experimento desenvolvido para a coleta de dados deste trabalho envolveu perguntas-guia associadas a figuras e foi dividido em três etapas subsequentes: (i) etapa de treinamento; (ii) etapa oral; (iii) etapa escrita. Na etapa (i), os participantes foram instruídos a memorizar o nome de algumas figuras para que se familiarizassem com os estímulos experimentais. Na etapa (ii), cada participante foi orientado a observar as figuras, ouvir uma pergunta-guia e respondê-la em voz alta. Nesta etapa, coletavam-se os dados de fala. Na etapa (iii), os participantes viram novamente os estímulos e escreveram a resposta da pergunta-guia em uma folha de papel.

As figuras e as perguntas foram organizadas no *Power Point* e houve aleatorização automática dos slides antes de cada coleta, visando a diminuir o risco de haver efeito (não controlado) da ordem das figuras nos dados. As perguntas-guia foram gravadas em áudio para que todos os participantes ouvissem o mesmo estímulo. O desenho experimental foi pensado para evitar interferências da forma oral da palavra na forma ortográfica. Utilizamos perguntas-guia para que os participantes falassem e escrevessem os estímulos com seu próprio conhecimento linguístico. A duração do experimento foi, em média, 25 minutos com cada participante. Todos os participantes receberam a mesma instrução.

Os estímulos experimentais utilizados na coleta dos dados foram selecionados no *Dicionário Ilustrado de Português* (Biderman, 2004). Esse dicionário foi escolhido porque, além de ser uma ferramenta voltada ao público infantil, as palavras que o constituem são de alta frequência e estão presentes em livros didáticos do Ensino Fundamental (Biderman, 2004, p. 8). O dicionário conta com cerca de 5.900 itens lexicais, dentre os quais foram selecionados os 24 estímulos experimentais deste trabalho, descritos no Quadro 2.

Quadro 2 - Estímulos experimentais.

<b>Categorias</b>	<b>Estímulos</b>
Pretônicas	tesoura – bezerro – besouro – vestido – peruca – bexiga- tomate – morcego - boneca - formiga – coruja – borracha
Postônicas finais	chave – alface – sorvete – lustre – cabide – peixe- garfo -laço – cavalo – caderno – gato – suco

Fonte: Elaborado pela autora.

Foram selecionadas palavras paroxítonas com <e> e <o> em posição pretônica e palavras paroxítonas terminadas com as letras <e> e <o> postônicas finais. Além disso, controlamos a classe gramatical das palavras, em busca de facilitar a escolha das figuras que constituíram os experimentos. Foram selecionados apenas substantivos concretos. Na categoria das vogais pretônicas, escolhemos seis palavras paroxítonas com a letra <e> pretônica e seis palavras paroxítonas com a letra <o> pretônica. Foram selecionadas apenas palavras trissílabas em que a sílaba pretônica é a primeira da palavra e a única antes da tônica, em busca de mantermos o padrão silábico e acentual. No caso das vogais postônicas finais, foram escolhidas seis palavras paroxítonas terminadas em <e> e seis palavras paroxítonas terminadas em <o>.

Os participantes da pesquisa foram crianças matriculadas no 1º, 3º, 5º, 7º e 9º anos escolares de uma escola municipal da região de Venda Nova em Belo Horizonte – MG. A escola municipal foi escolhida por trabalhar com todos os anos escolares do

Ensino Fundamental. Foram selecionados, aleatoriamente, dez alunos de cada ano escolar. Ao todo, 50 alunos participaram da pesquisa. Todos os participantes são naturais de Belo Horizonte – MG. A idade média dos participantes do 1º, 3º, 5º, 7º e 9º é, respectivamente, seis, oito, dez, 12 e 14 anos.

Cada participante produziu 24 dados referentes às vogais pretônicas e 24 dados referentes às vogais postônicas finais. Dos 48 dados produzidos por participante, 24 foram dados orais e 24 foram dados de escrita. No total, foram coletados 1.200 dados de escrita e 1.200 dados de fala (24 dados x 50 participantes). Após a filtragem de dados descartados por escrita ilegível ou por escolha de palavras diferentes dos estímulos, a amostra contou com 2286 dados, sendo 1.145 de fala (568 pretônicas + 577 postônicas) e 1.141 de escrita (565 pretônicas + 576 postônicas).

Os dados de escrita foram categorizados quanto aos erros e aos acertos relacionados às vogais pretônicas e postônicas finais. Erros ortográficos que saíram do escopo da pesquisa não foram quantificados<sup>2</sup>. A categorização dos erros e dos acertos possibilitou medir o número de vezes que os alunos de cada ano escolar acertaram a escrita das vogais pretônicas e postônicas finais. Os dados de fala foram codificados pelo som produzido nos contextos pretônico e postônico final. As variantes identificadas na fala foram divididas em dois grupos: (I) variantes próximas à forma gráfica (vogais médias); (II) demais variantes. O Quadro 3 sumariza os dados encontrados e as categorias elaboradas.

Quadro 3 - Variantes orais identificadas no conjunto de dados.

Categorias	Vogal pretônica	Vogal postônica final
(i) variantes próximas à forma gráfica	vogal média anterior vogal média posterior	vogal média anterior vogal média posterior
(ii) demais variantes	vogal alta anterior vogal alta posterior	vogal alta anterior vogal alta posterior

Fonte: Elaborado pelo autor.

A codificação ilustrada no Quadro 3 nos permitiu verificar se o ano escolar – e consequentemente o tempo de estudo da escrita – motiva o aumento de variantes próximas à forma gráfica (categoria (i)) na fala de alunos do Ensino Fundamental. Ou seja, a codificação nos possibilitou visualizar se, ao longo dos anos escolares, há aumento de variantes como laç[o], p[e]ruca, em detrimento de laç[u] e p[i]ruca.

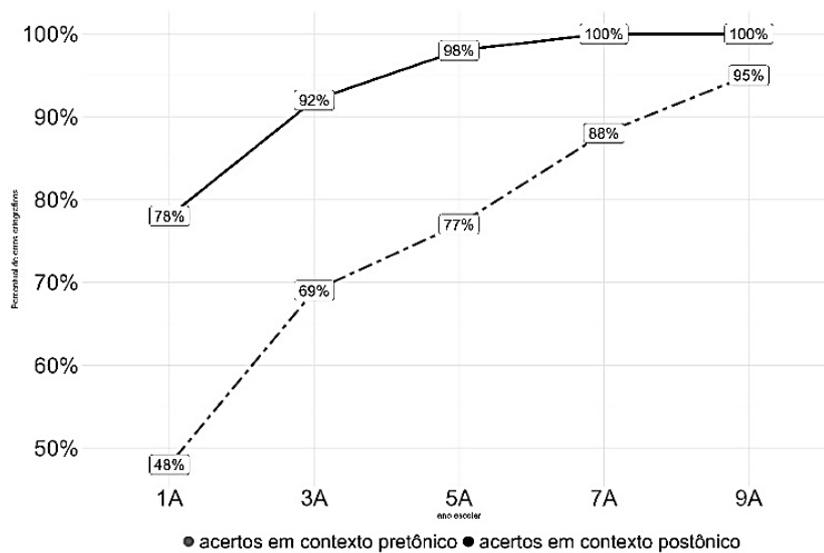
Os dados foram analisados por meio de métodos estatísticos descritivos e inferenciais. A análise estatística descritiva auxiliou na exploração e na visualização gráfica dos dados. Como as variáveis da pesquisa são categóricas, as descrições envolveram contagens e percentagens. Foram contabilizados os índices percentuais de variantes orais e de erros e acertos na escrita. Além da análise descritiva, ajustamos modelos lineares generalizados mistos ao conjunto de dados. Esses modelos avaliam, matematicamente, a relação entre duas ou mais variáveis (Godoy, 2019). No caso deste

<sup>2</sup> Vários tipos de erros, que fogem do escopo da pesquisa, foram encontrados nos dados. Por exemplo: <\*abi> (abrir); <\*aufasi> (alface); <\*bb> (beber); <\*bizoro, bizourro, bezolro,> (besouro); <\*mexiga> (bexiga); <\*bisero> (bezerro); <\*bonca> (boneca); <\*boraxa, \*borasa> (borracha); <\*cabite> (cabide); <\*cadeno> (caderno); <\*cacona; \*caukanha> (calcanhar); <\*cato> (cantor); <\*xavi> (chave); <\*coriji> (corrigir); <\*mucgo> (morego) etc.

trabalho, os modelos lineares foram utilizados para avaliar se a produção de variantes orais está relacionada ao contexto acentual e ao ano escolar (para maiores informações sobre as análises estatísticas, consulte Toledo, 2023). Toda a análise estatística foi feita por meio do Software R (R Core, 2022).

## 6 ANÁLISE E DISCUSSÃO DOS DADOS

Em primeiro lugar, é válido demonstrar como se deu o aprendizado das letras <e, o> em contexto pretônico e postônico final ao longo do Ensino Fundamental. Para tal, apresentamos, no Gráfico 1, a porcentagem de acertos – ou grafia conforme a norma ortográfica – por ano escolar e por contexto acentual. Vejamos.

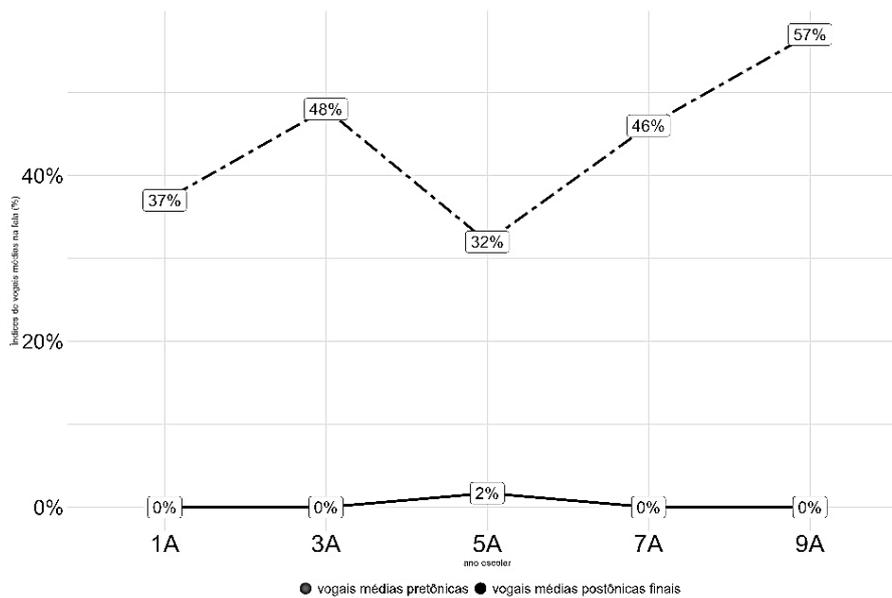


Fonte: Elaborado pela autora.

Gráfico 1 – Índices de acerto – escrita de <e, o> - por ano escolar em contextos pretônico e postônico final. A linha cinza pontilhada ilustra a porcentagem de acertos por ano escolar em contexto pretônico. A linha preta contínua se refere ao índice de acertos em contexto postônico final por ano escolar.

Como se pode ver, os alunos do Ensino Fundamental tendem a acertar a ortografia na medida em que progridem na escola. Em outros termos, os alunos passam a escrever as letras <e, o> corretamente em contextos pretônico e postônico final no decorrer da escolarização. Esses resultados eram esperados, haja vista que, ao longo da escolarização, os alunos recebem instruções que os auxiliam no desenvolvimento de habilidades ortográficas. Além disso, na medida em que avançam na escola, os aprendizes aumentam o contato com materiais gráficos, o que pode levá-los ao aprendizado implícito dos padrões de escrita da língua portuguesa (Treiman; Kessler, 2014).

Tendo em vista os resultados apresentados no Gráfico 1, passamos a avaliar se, com o desenvolvimento do saber ortográfico, os alunos alteram a maneira como falam as vogais pretônicas e postônicas finais. Temos como hipótese que o contexto acentual – pretônico ou postônico final – interfere no aumento de vogais médias na fala ao longo dos anos escolares. Vejamos o Gráfico 2.

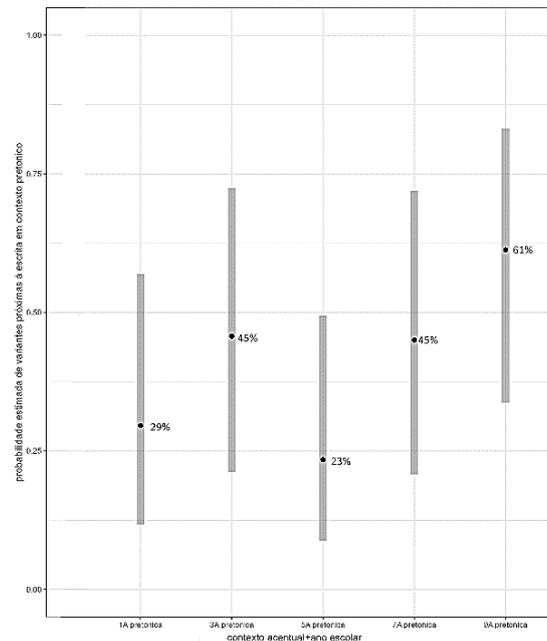


Fonte: Elaborado pela autora.

Gráfico 2 - A linha cinza pontilhada se refere aos índices de [e, o] pretônicas na fala por ano escolar. A linha preta contínua indica os índices de [e, o] postônicas finais na fala por ano escolar.

Observa-se, no Gráfico 2, o aumento de vogais médias ao longo dos anos escolares apenas em contexto pretônico. Em contexto postônico final, as vogais médias foram produzidas apenas no 5º ano e em índices extremamente baixos (2%). Um modelo linear generalizado misto<sup>3</sup> foi ajustado na amostra de dados de fala referente às vogais pretônicas e postônicas finais para testar a hipótese: o aprendizado das letras <e, o> motiva o aumento de vogais médias na fala a depender do contexto acentual: pretônico ou postônico final. A variável dependente do modelo foi binária: variantes próximas à forma gráfica/ demais variantes. O contexto acentual (pretônico/postônico final) e o ano escolar (1º, 3º, 5º, 7º e 9º) foram as duas variáveis independentes de efeito fixo do modelo. O item lexical e o participante foram as duas variáveis aleatórias ajustadas. Uma comparação de modelos aninhados mostrou que o melhor modelo ajustado continha a interação entre as variáveis *ano* e *contexto acentual* ( $\chi^2=9.8634$ ,  $p\text{-value}=0.04279$ ). Ou seja, a produção de vogais médias na fala é explicada pelo ano escolar em interação com o contexto acentual. Além disso, o modelo estimou probabilidades de ocorrência de vogais médias em contexto pretônico e postônico final em cada ano escolar. A probabilidade de ocorrer vogais médias em contexto postônico final foi 0% no 1º, 3º, 7º e 9º anos. Por isso, reportaremos apenas as probabilidades de ocorrência de vogais médias pretônicas por ano escolar. Observemos o Gráfico 3:

<sup>3</sup> Modelo <- glmer (TIPODEVARIANTE~ANO\*FENOMENO + (1|PARTICIPANTE) + (1|PALAVRA), dados.fala.vogais, family = binomial).



Fonte: Elaborado pela autora.

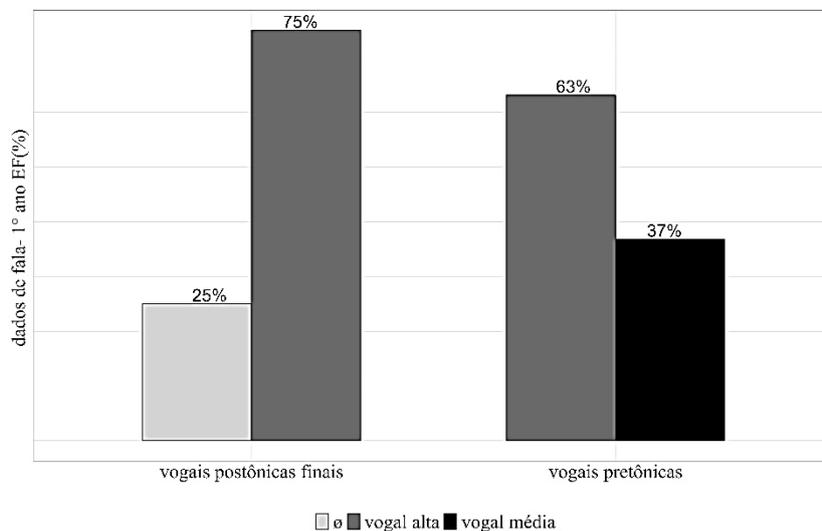
Gráfico 3 - Probabilidade de vogais médias pretônicas na fala por ano escolar.

Considerando-se que os dados se mantêm constantes, pode-se dizer que a probabilidade de ocorrer as vogais [e, o] pretônicas na fala oscila a cada ano escolar. Do 1º para o 3º ano, houve aumento da probabilidade de vogais médias na fala: de 29% para 45%. Do 3º para o 5º a probabilidade diminuiu: de 45% para 23%. Do 5º ano em diante a probabilidade voltou a aumentar: 23% < 45% < 61%. Esses resultados indicam, a princípio, que os aprendizes de diferentes anos escolares produzem diferentes índices de vogais [e, o] na fala. Há de se ressaltar, contudo, que a grande amplitude dos intervalos de confiança não nos permite garantir que a probabilidade de produção de vogais médias difere significativamente em todos os anos escolares. Fizemos uma análise *post-hoc* para comparar par a par os anos escolares. Os resultados dessa análise confirmaram que há diferenças significativas entre as amostras das vogais pretônicas do 5º e do 9º anos ( $p\text{-value} = 0.0390$ ). Ou seja, podemos confirmar que os alunos do 5º ano falam menos vogais médias pretônicas do que os alunos do 9º ano. No caso das vogais postônicas finais, não houve diferenças significativas entre nenhum ano escolar.

De forma geral, os resultados confirmam a hipótese de que o aumento de vogais médias na fala ao longo dos anos escolares depende do contexto accentual: pretônico ou postônico final. Especificamente, os dados indicam que o aumento da escolaridade – e conseqüentemente o aumento do contato com palavras grafadas com as letras <e, o> – motiva o aumento de vogais médias na fala em contexto pretônico, mas não em contexto postônico final. Por exemplo: os alunos aumentam a produção de formas como [me'nino] e [bo'neca], mas continuam falando, sistematicamente, ['pêti] e ['patu]. Resta-nos, pois, perguntar: *por que os aprendizes aumentam a pronúncia de vogais médias pretônicas ao longo da escolarização, mas não fazem o mesmo em contexto postônico final?* Com o apoio dos Modelos de Exemplares, formulamos a seguinte resposta à pergunta: quando os indivíduos aprendem a escrever, há reorganização das representações mentais para que o conhecimento ortográfico se acomode ao sistema linguístico. Nesse processo, os aprendizes podem alterar a forma como falam. No

entanto, as categorias vocálicas são múltiplas e probabilísticas e, por isso, a escrita ortográfica opera de maneira distinta em cada uma delas. Em outros termos, os efeitos da escrita na fala são diferentes em contexto pretônico e postônico final, porque as categorias abstratas dessas vogais são diferentes. Em vez de serem únicas e discretas (fonemas), as categorias, nos Modelos de Exemplos, são múltiplas por armazenarem informações fonéticas, lexicais, contextuais, semânticas, socioculturais etc. Assim sendo, é coerente afirmar que as vogais pretônicas e postônicas estão armazenadas em diferentes categorias abstratas, haja vista que ocorrem em diferentes contextos e estão envolvidas em diferentes fenômenos fonológicos no português brasileiro

Para exemplificar a natureza múltipla e probabilística das categorias vocálicas, apresentamos, a seguir, os dados de fala dos alunos do 1º ano do Ensino Fundamental, organizados por contexto – pretônico ou postônico final.

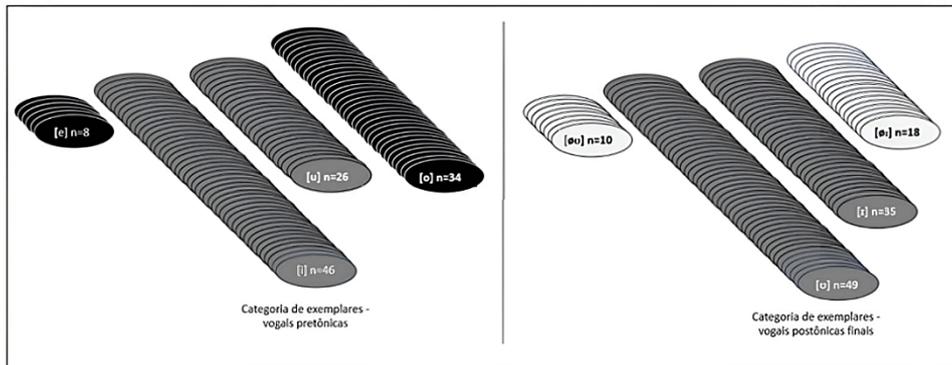


Fonte: Elaborado pela autora.

Gráfico 4 - Variação sonora em contextos postônico final e pretônico.

As colunas à esquerda do Gráfico 4 se referem às variantes orais produzidas em contexto pretônico por alunos do 1º ano do EF. Como se pode ver, esses alunos produziram 63% de vogais altas e 37% de vogais médias em contexto pretônico. As colunas à direita, por sua vez, se referem às variantes produzidas em contexto postônico final. Nesse contexto, os alunos apagaram as vogais em 25% dos dados e produziram vogais altas em 75%. O que os dados nos mostram, em síntese, é que os alunos do 1º ano do EF produzem variantes orais distintas em contextos pretônico e postônico final.

Partindo do pressuposto de que as unidades linguísticas armazenadas na memória humana são unidades atestadas no uso real da língua (Bybee, 2001), pode-se dizer que os dados expostos no Gráfico 4 refletem parte do conhecimento fonológico dos aprendizes belorizontinos. Em outras palavras, os dados de fala dos alunos do 1º ano do EF evidenciam que as categorias fonológicas das vogais pretônicas e postônicas finais armazenam exemplares distintos. O diagrama representado na Figura 1 ilustra a multiplicidade das categorias fonológicas, a partir dos dados de fala dos alunos do 1º ano do EF.



Fonte: Elaborado pela autora.

Figura 1 - Diagrama com nuvens de exemplares. À esquerda, tem-se a categoria de exemplares das vogais pretônicas; à direita, das vogais postônicas finais.

Os ovais à esquerda do diagrama ilustram as nuvens de exemplares referentes às vogais pretônicas. Os ovais à direita mostram os exemplares registrados na categoria das vogais postônicas finais. A quantidade de ovais é igual à quantidade total de variantes orais produzida no conjunto de dados dos alunos do 1º ano do EF. Em preto, tem-se as nuvens de exemplares das vogais médias. Observe que as vogais médias foram armazenadas apenas na categoria das vogais pretônicas – [e] (n=8); [o] (n=34). Em cinza escuro, estão as nuvens de exemplares das vogais altas. Em contexto pretônico, há mais exemplares de vogais altas anteriores (n=46) do que posteriores (n=26). Em contexto postônico final, por outro lado, há mais vogais altas posteriores (n=49) do que anteriores (n=35). As vogais anteriores postônicas finais estão mais envolvidas em cancelamento (n=18) do que as vogais posteriores postônicas finais (n=10). De modo geral, o diagrama da Figura 1 foi elaborado a fim de demonstrar que as categorias das vogais pretônicas e postônicas finais armazenam múltiplos e distintos exemplares.

É necessário ressaltar ainda o porquê de se apresentar os dados dos alunos do 1º ano do EF. Dentre os participantes desta pesquisa, os alunos do 1º ano são os que têm menos tempo de contato com a escrita em ambiente escolar<sup>4</sup>. Isto é, no 1º ano do EF, os alunos estão sendo alfabetizados. Pode-se dizer, então, que o modo de falar desses alunos é menos afetado pela escrita do que o dos demais alunos participantes desta pesquisa. Sabendo-se disso, podemos pensar que, no conjunto de dados aqui analisado, o 1º ano pode ser tomado como o estágio inicial de aprendizado formal da escrita. Ou seja, a amostra de dados dos alunos do 1º ano nos traz evidências sobre o conhecimento linguístico prévio à consolidação do saber ortográfico. Pudemos ver, ao longo da análise, que, desde o primeiro ano, os alunos produziram vogais médias pretônicas. Esse fato pode explicar o porquê de haver aumento dessas vogais no decorrer do aprendizado da escrita. Como já foi mencionado, argumentamos, neste artigo, que, ao longo do aprendizado da escrita, há reorganização das representações mentais para que o conhecimento ortográfico se acomode ao sistema linguístico. Mais

<sup>4</sup> Alguns alunos do 1º ano podem ter contato com a escrita fora da escola desde tenra idade. Assim, sabemos que não é possível afirmar que as crianças do 1º ano têm o primeiro contato com escrita apenas quando entram na escola. No entanto, é coerente dizer que os alunos do 1º ano têm menos contato com a escrita se comparado com os alunos dos demais anos escolares avaliados nesta pesquisa – 3º, 5º, 7º e 9º anos.

do que isso, argumentamos que as formas ortográficas recém aprendidas podem fortalecer determinadas nuvens de exemplares.

Vale ressaltar, todavia, que a força da escrita na representação mental varia a depender da robustez dos exemplares. Como foi visto no diagrama (Figura 1), os alunos, já no 1º ano do EF, têm uma grande quantidade de vogais médias pretônicas registradas na memória. Esses exemplares podem se fortalecer ainda mais com o aprendizado das letras <e, o>. Isto é, a representação de palavras como [me'ninʊ] e [bo'nekə] pode se fortalecer quando associada às formas ortográficas que as representam: <menino> e <boneca>. O fortalecimento da representação fonológica explica o aumento dessas variantes na fala. Já no caso das vogais postônicas finais, as vogais médias [e, o] têm representação muito menos robusta do que as vogais altas [ɪ, ʊ]. Isso porque, conforme evidenciam os dados, as vogais médias postônicas finais ocorrem em menor frequência do que as vogais altas desde o 1º ano do EF. O aprendizado das letras <e, o> pode até fortalecer a representação das vogais médias postônicas finais. Por exemplo: a forma ortográfica <pente> pode motivar o fortalecimento da representação fonológica [pête]. Porém, as vogais médias são tão raras em posição postônica final que, dificilmente, as formas ortográficas terão força suficiente para moldar a representação a ponto de que [e, o] passem a ser frequentes na oralidade. Em outros termos, as letras <e, o> podem até fortalecer a representação de [e, o] postônicos finais, mas não de forma suficiente para que os índices de vogais médias postônicas finais se igualem ou ultrapassem os índices de vogais altas postônicas finais na fala dos aprendizes.

Em resumo, queremos argumentar que as letras <e, o> são acomodadas no sistema linguístico e, mais do que isso, podem fortalecer a representação e o uso de variantes sonoras. No entanto, a força da escrita é maior em padrões que já são produtivos na língua: como [e, o] pretônicos. Em caso de padrões não produtivos – como [e, o] postônico final – a escrita não tem força suficiente para alterar as representações e o uso linguístico. Este artigo avança, então, ao argumentar que a fala e a escrita podem estabelecer múltiplas relações. Em vez de se limitar à ideia unirepresentacional de que letras representam fonemas, propomos que uma mesma letra – <e, o> – pode estar relacionada a categorias linguísticas múltiplas e probabilísticas.

## 7 CONSIDERAÇÕES FINAIS

Este artigo investigou a influência do aprendizado das letras <e, o> na produção oral de vogais médias e altas pretônicas e postônicas finais. Os dados indicaram que tanto as vogais pretônicas quanto as vogais postônicas finais estão envolvidas em fenômenos de variação sonora. No contexto pretônico, a variação se deu entre as vogais médias e altas. Exemplo: [mi'ninʊ] ~ [me'ninʊ]. No contexto postônico final, a variação se deu entre vogais altas e canceladas. Exemplo: ['patʊ] ~ ['pat]. Ou seja, os resultados evidenciam que as vogais pretônicas e postônicas finais se manifestam de forma distinta no falar de Belo Horizonte -MG.

Observamos, sobretudo, como se deu a produção de vogais médias pretônicas e postônicas finais, visando a verificar se o aprendizado das letras <e, o>, ao longo dos anos escolares, motivou o aumento de vogais médias na oralidade. Houve aumento de vogais médias pretônicas, sobretudo, na fala de alunos do 9º ano. Por outro lado,

não houve aumento de vogais médias postônicas finais ao longo dos anos escolares. Portanto, confirmou-se a hipótese de que o aumento de vogais médias na fala ao longo dos anos escolares depende do contexto acentual: pretônico ou postônico final.

Apoiando-se nas premissas dos Modelos de Exemplares, esses resultados foram interpretados como evidências de que as formas ortográficas podem motivar a reorganização do conhecimento linguístico. Porém, o efeito da escrita na representação mental pode ser mais ou menos evidente a depender da categoria linguística, uma vez que estas são múltiplas, dinâmicas e probabilísticas. Argumentamos que padrões sonoros já produtivos na língua – como as vogais médias pretônicas – podem se tornar ainda mais produtivos no sistema linguístico após o aprendizado da escrita. Por outro lado, em padrões de baixa produtividade – como [e, o] postônicos finais –, a forma ortográfica não tem força suficiente para reorganizar o sistema a ponto de aumentar significativamente a produtividade do padrão sonoro.

Vale ressaltar, por fim, que este trabalho se limitou a investigar dados transversais de fala e de escrita. Esse tipo de análise nos permitiu comparar amostras de dados de diferentes alunos, mas nos impossibilitou de avaliar possíveis alterações no comportamento oral e gráfico de um mesmo indivíduo. Isto é, não nos foi possível afirmar que um mesmo indivíduo altera a fala após o aprendizado da escrita. Análises futuras podem se utilizar de metodologia longitudinal para investigar se uma criança pré-alfabetizada, por exemplo, altera sua fala ao longo da alfabetização. Esse tipo de análise poderá confirmar, com mais ênfase, os efeitos da escrita na organização do conhecimento linguístico.

Esperamos, de modo geral, que a presente pesquisa possa promover reflexões acerca das diferentes relações estabelecidas entre a fala e a escrita. Além disso, esperamos que este trabalho consiga destacar a relevância em se considerar a multiplicidade e a dinamicidade da gramática para expressar as relações entre oralidade e escrita.

## REFERÊNCIAS

Adamoli MA. Um estudo sobre o estatuto fonológico dos ditongos variáveis [aj] e [ej] do PB a partir de dados orais e ortográficos produzidos por crianças de séries iniciais [tese]. Pelotas: Faculdade de Educação, Universidade Federal de Pelotas; 2012.

Alvarenga D, et al. Da forma sonora da fala à forma gráfica da escrita uma análise linguística do processo de alfabetização. *Caderno de Estudos Linguísticos*. 1989;16:5-30.

Assis A. A emergência de consoantes finais no português brasileiro na microrregião de Araguaína/Tocantins [tese]. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2017.

Biderman M. *Dicionário ilustrado de português*. São Paulo: Ática; 2004.

Bisol L. Harmonização vocálica: uma regra variável [tese]. Rio de Janeiro: Faculdade de Letras, Universidade Federal do Rio de Janeiro; 1981.

Bisol L. A neutralização das átonas. *Revista Letras*. 2003. 61(Nº esp):273-283.

Bod R, Hay J, Jannedy S. Introduction. In: Bod R, Hay J, Jannedy S, editores. *Probability theory in Linguistics*. Cambridge: MIT Press; 2003. p. 1-10.

Bybee J. *Phonology and language use*. Cambridge: Cambridge University Press; 2001.

Bybee J. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*. 2002;14(3):261-290.

- Chaves I. Panorama dos estudos das vogais pretônicas no português do Brasil: meta-análise das pesquisas desenvolvidas de 1980 a 2012 [tese]. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2013.
- Chevrot J-P, Beaud L, Varga R. Developmental data on a French sociolinguistic variable: post-consonantal word-final /R/. *Language Variation and Change*. 2000;12(3):295-319.
- Cristófaros-Silva T. Fonética e fonologia do português: roteiro de estudos e guia de exercícios. 10.<sup>a</sup> ed. São Paulo: Contexto; 2013.
- Cristófaros-Silva T. Dicionário de fonética e fonologia. São Paulo: Contexto; 2017.
- Cristófaros-Silva T, Gomes C. Representações múltiplas e organização do componente fonológico. *Fórum Linguístico*. 2004;4(1):147-177.
- Cristófaros-Silva T, Greco A. Representações fonológicas: contribuições da oralidade e da escrita. *Letras de Hoje*. 2010;45(3):87-93.
- Cristófaros-Silva T, Guimarães D. A aquisição da linguagem falada e escrita: o papel da consciência linguística. *Letras de Hoje*. 2013;48(3):316-323.
- Cristófaros-Silva T, Vieira MJ. Redução vocálica em postônica final. *Revista da Abralin*. 2015;14(1):379-406.
- Crystal D. Dicionário de linguística e fonética. Dias MCP, tradutora e adaptadora. Rio de Janeiro: Jorge Zahar; 1988. (Vol. 1).
- Dias E, Seara I. Redução e apagamento de vogais átonas finais na fala de crianças e adultos de Florianópolis: uma análise acústica. *Letrônica*. 2013;6(1):71-93.
- Fontes-Martins R. A organização do componente fonológico e o comportamento do indivíduo [tese]. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2007.
- Foulkes P, Docherty G. The social life of phonetics and phonology. *Journal of Phonetics*. 2006;34(4):409-438.
- Godoy M. Introdução aos modelos lineares mistos para os estudos da linguagem. *PsyArXiv*; 2019. [citado 20 abr. 2024]. Disponível em: <https://doi.org/10.17605/OSF.IO/9T8UR>.
- Gombert J. Atividades metalinguística e aquisição da leitura. In: Maluf MR, organizador. *Metalinguagem e aquisição da escrita*. São Paulo: Casa do Psicólogo; 2003. p. 19-64.
- Huback A. Cancelamento do (r) final em nominais na cidade de Belo Horizonte: uma abordagem difusionista [dissertação]. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2003.
- Johnson K. Speech perception without speaker normalization: an exemplar model. In: Johnson K, Mullenix JW, editores. *Talker variability in speech processing*. San Diego: Academic Press; 1997. p. 145-165.
- Kato M. No mundo da escrita: uma perspectiva psicolinguística. 3.<sup>a</sup> ed. São Paulo: Ática; 1990.
- Lemle M. Guia teórico do alfabetizador. São Paulo: Editora ática; 1987.
- Lemos F. Interferência da oralidade na escrita: o caso do registro ortográfico de e, i, o, u átonos. *Letras de Hoje*. 2001;36(3):465-471.
- Magalhães J, Campos R. As vogais médias pretônicas em Minas Gerais e em São Paulo: o Rio Grande como fronteira linguística. *Revista do GELNE*. 2022;24(2):152-167.
- Man J. A história do alfabeto. Zonenschain E, tradutora. Rio de Janeiro: Ediouro; 2002.
- Mattingly I. Linguistic awareness and orthographic form. In: Frost R, Katz L, editores. *Orthography, phonology, morphology, and meaning*. Amsterdam: Elsevier Science Publishers; 1992. p. 11-26.
- Miranda AR. Um estudo sobre a aquisição ortográfica das vogais do português. *Anais da ANPed Sul*; 2006; Santa Maria.
- Miranda AR. A aquisição ortográfica das vogais do português: relações com a fonologia e a morfologia. *Letras*. 2008;(36):151-168.

- Miranda AR. Um estudo sobre o erro ortográfico. In: Heining O, Fronza CA, organizadores. Diálogos entre linguística e educação. Blumenau: EDIFURB; 2010. p. 141-16. (Vol. 1).
- Morais J. A arte de ler. São Paulo: UESP; 1996.
- Morais J, et al. Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*. 1979;7(4):323-331.
- Nagy WE, Anderson RC. Metalinguistic awareness and literacy acquisition in different languages. In: Wagner DA, Venezky RC, Street B, editores. *Literacy: an international handbook*. Boulder: Westview Press; 1999. p. 155-160.
- Nóbrega MJ. Ortografia. São Paulo: Melhoramentos; 2013.
- Oliveira M. Aspectos da difusão lexical. *Revista de Estudos da Linguagem*. 1992;1(1):31-41.
- Oslo D. Literacy and language development. In: Wagner DA, Venezky RC, Street BV, editores. *Literacy: an international handbook*. Boulder: Westview Press; 1999. p. 132-136.
- Paiva M. O percurso da monotongação [ey]: observações no tempo real. In: Paiva MC, Duarte E, organizadores. *Mudança linguística em tempo real*. Rio de Janeiro: FAPERJ/Contra Capa; 2003. p. 31-46.
- Pierrehumbert J. Exemplar dynamics: word frequency, lenition and contrast. In: Bybee J, Hopper PJ, organizadores. *Typological studies in language*. Amsterdam: John Benjamins Publishing Company; 2001. p. 137. (Vol. 45). [citado 10 jul. 2021]. Disponível em: <https://benjamins.com/catalog/tsl.45.08pie>.
- Pierrehumbert J. Probabilistic theories of phonology. In: Bod R, Hay JB, Jannedy S, editores. *Probability theory in linguistics*. Cambridge: MIT Press; 2003. p. 177-228.
- Port R. How are words stored in memory? Beyond phones and phonemes. *New Ideas in Psychology*. 2007;25(2):143-170.
- R Core Team. R: a language and environment for statistical computing [programa de computador]. Viena: R Foundation for Statistical Computing; 2022. [citado 06 mai. 2024]. Disponível em: <https://www.r-project.org/>.
- Seymour PHK. Early reading development in European orthographies. In: Snowling MJ, Hulme C, editores. *The science of reading: a handbook*. Oxford: Blackwell Publishing; 2005. p. 296-315.
- Schiwindt L, et al. A influência da variável escolaridade em fenômenos fonológicos variáveis: efeitos retroalimentares da escrita. *RevEL: Revista Virtual de Estudos da Linguagem*. 2007;5(9):1-12.
- Soares M. Alfabetização e literatura. *Guia da alfabetização*. Educação (São Paulo). 2010;2:12-29.
- Soares M. Alfabetização: a questão dos métodos. São Paulo: Contexto; 2018.
- Souza R. Redução de vogais altas pretônicas no português de Belo Horizonte: uma abordagem baseada na gradiência [dissertação]. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2012.
- Tenani L, Reis M. “E viveram felizes para sempre”: análise de grafias não convencionais de vogais pretônicas. *Verba Volant*. 2011;2(1):22-43.
- Treiman R, Kessler B. How children learn to write words. New York: Oxford University Press; 2014.
- Toledo C. Relações múltiplas entre oralidade e escrita [tese]. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2023.
- Veloso J. A língua escrita e a escrita da língua. Algumas considerações sobre transparência e opacidade fonêmicas na escrita do português e outras questões. *Da Investigação às Práticas*. 2005;6(1):49-69.
- Viegas M. O acento de vogais médias pretônicas e os itens lexicais. *Revista de Estudos da Linguagem*. 1995;3(2):101-122.



## O processamento lexical do acento em pseudopalavras a partir da leitura de falantes do português brasileiro

### *The lexical processing of stress in pseudowords based on the reading of Brazilian Portuguese speakers*

Aline de Lima Benevides\*

Faculdade de Tecnologia do Estado de São Paulo, São Paulo, SP, Brasil

**Resumo:** Este trabalho investigou, a partir de um estudo experimental, o papel da similaridade fonológica, da frequência de ocorrência da palavra base e da sílaba final na atribuição do acento primário em pseudopalavras a partir da produção de falantes nativos do português brasileiro. A métrica de criação das pseudopalavras baseou-se no estudo de Protopapas et al. (2006). As palavras a partir das quais as pseudopalavras foram criadas tinham alta ou baixa frequência de ocorrência, três sílabas de extensão. Contrastaram-se quatro grupos experimentais: i. similar de alta frequência; ii. dissimilar de alta frequência; iii. similar de baixa frequência; e iv. dissimilar de baixa frequência. Hipotetizou-se que pseudopalavras similares de alta frequência desencadeassem mais processos analógicos do que pseudopalavras dissimilares. Participaram do estudo 34 falantes, em uma tarefa de leitura. Os resultados demonstraram que a semelhança entre palavras reais e pseudopalavras desencadeia processos analógicos com a extensão do padrão acentual e que a frequência de ocorrência da palavra alvo e a sílaba final também são relevantes. Tais resultados trazem indícios de que o acento é armazenado juntamente com a palavra (isto é, é lexical) e de que os níveis segmental, silábico e acentual interagem na determinação do local em que o acento incide.

**Palavras-chave:** Acento. Pseudopalavras. Português Brasileiro.

**Abstract:** This work investigated, based on an experimental study, the role of phonological similarity, the frequency of occurrence of the base word and the final syllable in the attribution of primary stress in pseudowords based on the production of native speakers of Brazilian Portuguese. The pseudoword creation metric was based on the study by Protopapas et al. (2006). The words from which the pseudowords were created had a high or low frequency of occurrence, three syllables in length. Four experimental groups were compared: i. similar high frequency; ii. high frequency dissimilar; iii. similar low frequency; and iv. low frequency dissimilar. It was hypothesized that similar high-frequency pseudowords would trigger more analogical processes than dissimilar pseudowords. 34 speakers participated in the study, in a reading task. The results demonstrated that the similarity between real words and pseudowords triggers analogical processes with the extension of the stress pattern and that the frequency of occurrence of the target word and the final syllable are also relevant. Such results provide evidence that the stress is stored together with the word (that is, it is lexical) and that the segmental, syllabic and stress levels interact in determining the place where the stress falls.

**Keywords:** Stress. Pseudowords. Brazilian Portuguese.

---

\* Doutorado em Letras pela Universidade de São Paulo; professora na educação básica e na FATEC; São Paulo, SP; [benevides.aline12@gmail.com](mailto:benevides.aline12@gmail.com)

## 1 INTRODUÇÃO

Mais de meio século após a primeira hipótese sobre o acento em português brasileiro (Câmara Jr., 1970), este fenômeno continua alvo de inúmeras análises e reanálises, visto que, no nosso entendimento, não há, ainda, uma proposta que consiga compreender todos os padrões acentuais sem se valer de estratégias de excepcionalidade. Elas, em geral, consideram aspectos prosódicos e/ou morfológicos (cf. Mateus, 1983; Bisol, 1994; Lee, 1995, 2007; Massini-Cagliari, 1999; Magalhães, 2004; Wetzels, 2007), além de justificar padrões ditos irregulares como extramétricos e/ou marcados, por exemplo.

Atualmente, a proposta que melhor justifica os padrões encontrados na língua consiste na de Cantoni (2013), uma das mais recentes análises do acento em Português Brasileiro (doravante PB). A autora assume dois mecanismos centrais: especificação lexical e determinação morfológica. Estes são explicados diacronicamente. O latim, língua-mãe do português, caracterizava-se como uma língua de acento fixo, visto que era determinado pela quantidade silábica: sílabas penúltimas pesadas atraíam o acento, do contrário, o acento recuava para a sílaba precedente (antepenúltima). Por o português ter herdado grande parte do léxico do latim, há, ainda em seu léxico, resquícios dessa forte correlação entre peso silábico e padrão acentual. Entretanto, diante do fato de que a sensibilidade à quantidade silábica foi perdida na passagem do latim ao português e de que inúmeros processos fonológicos atingiram a língua nesse período, a previsibilidade do acento se perdeu e a regularidade dos padrões morfológicos se estabeleceu como o novo atrator da língua, determinando o local em que o acento incide. Com isso, a morfologia passa a regular o padrão acentual de palavras que apresentam morfemas de derivação ou de flexão. Os demais vocábulos são compreendidos como especificados lexicalmente, já que teriam herdado a tonicidade da língua da qual provém.

O ponto central da análise de Cantoni (2013), que nos faz assumi-la como ponto de partida do presente trabalho, são os mecanismos que regulam a construção das representações gramaticais. Para ela, os vocábulos, ao serem usados, têm seus padrões abstraídos, categorizados e armazenados em suas representações mentais, de forma que eles podem ser replicados a novas palavras - conforme previstos pelos Modelos baseados no Uso (Bybee, 2001, 2006, 2010) e pela Teoria dos Exemplares (Kemmer; Barlow, 1999; Langacker, 1999, 2008, 2013; Bybee, 2010; Pierrehumbert, 2001). Esses modelos também preveem que a frequência de tipo e a de ocorrência são mediadoras desse processo, já que a frequência de tipo é responsável pela fortificação e pela replicação de padrões recorrentes (silábicos, segmentais, acentuais etc.), ao passo que a frequência de ocorrência garante a estabilização e a não regularização de vocábulos irregulares.

Todos esses mecanismos se mostram relevantes para a presente pesquisa, tendo em vista que testamos, à luz do experimento realizado por Protopapas et al. (2006) na língua grega, em que medida a semelhança entre palavras reais e pseudopalavras pode desencadear processos analógicos, fazendo com que o padrão acentual da palavra real seja estendido para o novo vocábulo - por exemplo, investiga-se, se uma pseudopalavra como *catater*, ao ser associada com a palavra real *caráter*, em decorrência de suas semelhanças fonológicas, receberá também o padrão acentual paroxítono (*catáter*). Em outros termos, avaliamos a hipótese de que a produção de

pseudopalavras<sup>1</sup> pode resgatar e ativar padrões métricos e segmentais baseados nas palavras do PB armazenadas no léxico mental dos falantes, a partir de processos analógicos entre palavras reais e pseudopalavras. Para realizar tal análise, três variáveis são analisadas: similaridade fonológica, frequência de ocorrência do item lexical e sílaba final. Através da similaridade fonológica, investigamos em que medida pseudopalavras que são semelhantes a palavras reais podem ter o seu padrão acentual estendido para pseudopalavras. Ou seja, a questão que está posta é: ao processarmos um vocábulo novo (pseudopalavra), associamo-lo a uma palavra existente e resgatamos o seu acento? Quanto à frequência dos itens lexicais, analisamos se a frequência de ocorrência do item lexical a partir do qual a pseudopalavra é criada pode mediar o acesso lexical, promovendo mais produções com semelhante padrão acentual em palavras de alta frequência do que as de baixa frequência. A sílaba final, por fim, permite-nos tanto controlar o *template* das pseudopalavras, como também testar se a composição da sílaba final pode ou não ser atratora do acento. Isso porque propostas métricas, como a de Bisol (1994), propõem que sílabas pesadas finais são atradoras do acento em português.

É preciso destacar, a priori, que esta pesquisa não tem o intuito de advogar a favor de que um único fator determina o acento em PB; ao contrário disso, partimos da hipótese de Cantoni (2013) de que o acento é lexical, de que padrões acentuais foram abstraídos e podem ser reproduzidos em novos vocábulos. Buscamos, portanto, mapear quais são os fatores que contribuem para que tal processo ocorra. Esses fatores, em nossa perspectiva teórica, são múltiplos. Este trabalho se dedica a avaliar apenas três: similaridade fonológica, frequência de ocorrência do item lexical e sílaba final, a partir de modelos de processamento de palavra (Coltheart; Rastle, 1994; Rastle; Coltheart, 2000; Coltheart, 2005), dos Modelos de Uso (Bybee, 2001, 2006, 2010) e de Exemplos (Kemmer; Barlow, 1999; Langacker, 1999, 2008, 2013; Bybee, 2010; Pierrehumbert, 2001).

## 2 METODOLOGIA

### 2.1 Participantes

Participaram deste estudo 34 falantes nativos do português brasileiro, com idade entre 18 e 60 anos, de ambos os sexos e com nível de escolaridade de Ensino Fundamental completo a Mestrado. Nenhum dos participantes relatou qualquer problema cognitivo, auditivo ou fonoarticulatório<sup>2</sup>.

### 2.2 Estímulos

Todas as pseudopalavras apresentam três sílabas com os seguintes *templates*: CV-CV-CV ou CV-CV-CVC. Selecionaram-se essas estruturas por serem as mais frequentes do português brasileiro (Benevides; Guide, 2017) e, ainda, por permitirem

---

<sup>1</sup> Pseudopalavras são palavras inventadas que respeitam a fonotática da língua, mas não possuem significado.

<sup>2</sup> Esta pesquisa foi registrada e aprovada pelo Comitê de Ética em Pesquisa com Seres Humanos da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo (EACH/USP), sob o processo nº 3.332.409. Todos os participantes consentiram que o estudo fosse gravado e utilizado em pesquisas acadêmicas.

que todas as tonicidades (oxítone, paroxítone e proparoxítone) pudessem ser empregadas pelos participantes. Além disso, esses *templates* permitem avaliar em que medida a composição da sílaba final, com ou sem coda, pode determinar o local em que o acento incide. Investiga-se, com esses *templates*, possíveis efeitos da sílaba final.

Tais estruturas foram preenchidas por material segmental a partir da métrica de similaridade fonológica proposta por Protopapas et al. (2006). Segundo os autores, palavras similares são aquelas que se diferenciam minimamente, numa escala de 1 a 3, considerando, para isso, modificação de vozeamento, ponto e modo de articulação - um ponto para cada mudança. A título de exemplo, a pseudopalavra *tarater* apresenta similaridade fonológica de um nível com a palavra *caráter*, pois a mudança de /k/ para /t/ acarreta a modificação apenas de ponto de articulação. Palavras dissimilares, por sua vez, são aquelas que se diferenciam de 5 a 10 níveis - novamente, considerando vozeamento, ponto e modo de articulação. A palavra *caráter* é dissimilar a *gatafer*, uma vez que há mudança no vozeamento de /k/ para /g/; troca de /r/ para /t/ com modificação de vozeamento e de ponto de articulação; e, por fim, alteração de /t/ para /f/, totalizando cinco mudanças. Há, portanto, dois níveis da variável similaridade: similar e não similar.

A frequência de ocorrência do item lexical a partir do qual a pseudopalavra foi criada consistiu no terceiro fator de investigação. Foram caracterizadas como palavras de alta frequência aquelas que apresentavam frequência igual ou superior a 100 ocorrências por milhão e palavras de baixa frequência aquelas cuja frequência era igual ou inferior a 2 ocorrências por milhão. A frequência de ocorrência foi retirada do Corpus ABG (Benevides; Guide, 2017). Para esta variável, dois níveis foram investigados: alta e baixa frequência.

As variáveis similaridade fonológica e frequência de ocorrência foram cruzadas, formando quatro grupos de investigação: i. pseudopalavras maximamente similares de alta frequência; ii. pseudopalavras minimamente similares de alta frequência; iii. pseudopalavras maximamente similares de baixa frequência; e iv. pseudopalavras minimamente similares de baixa frequência, conforme exposto no Quadro 1<sup>3</sup>.

Quadro 1 - Descrição dos grupos de pseudopalavras por similaridade e frequência com exemplos.

	Similaridade	Frequência	Exemplos
<b>Grupo 1</b>	Alta	Alta	Figura ( <i>dígura, fidura, figuba</i> ) Caráter ( <i>darater, cacater, caraper</i> )
<b>Grupo 2</b>	Baixa	Alta	Código ( <i>nojilo, zofisu, bocito</i> ) Militar ( <i>pitudar, dipinar, sipivar</i> )
<b>Grupo 3</b>	Alta	Baixa	Pateta ( <i>lateta, pafeta, pateda</i> ) Cabaré ( <i>nabare, camare, cabate</i> )
<b>Grupo 4</b>	Baixa	Baixa	Pêssego ( <i>lemifo, demefo, zedeno</i> ) Funeral ( <i>mupeval, lusipal, puvebal</i> )

Fonte: adaptado de Benevides (2022, p. 128).

<sup>3</sup> Confira todos as palavras reais e as pseudopalavras no Apêndice.

A fim de evitar possíveis enviesamentos, restringimos o uso de consoantes palatais (/ɲ/ e /ʎ/), da fricativa /x/ e das pós-alveolares (/ʃ/ e /ʒ/) em posição de ataque da última e/ou da penúltima sílabas, visto que, segundo Câmara Jr. (1953), Cristófaros-Silva (1995) e Wetzels (2000), esses segmentos impedem a emergência do acento proparoxítono - o que é contrário ao objetivo deste estudo, que é permitir que todas as tonicidades possam emergir.

### 2.3 Validação da métrica de (dis)similaridade

Para garantir que os estímulos criados fossem, de fato, considerados (dis)similares a uma palavra real, todas as pseudopalavras criadas foram submetidas a um teste de validação. Este consistiu na exposição de todos os vocábulos, por meio do envio de um formulário *on-line*, a um grupo de 10 falantes nativos do PB, que foram instruídos a indicar para cada pseudopalavra uma palavra real que se assemelhasse - isto é, a primeira que viesse em mente. O índice determinado para que o estímulo fosse validado como similar foi de 80% - ou seja, dos 10 participantes, 8 deveriam indicar a mesma palavra que utilizamos como base para a criação da pseudopalavra; para os dissimilares, dos 10 participantes, no máximo, 2 indivíduos podiam indicar uma mesma palavra, independente se era a palavra base ou qualquer outro vocábulo.

A taxa média de validação foi de 72,2% das palavras. Somente as pseudopalavras validadas por esse teste foram incluídas na presente análise, de forma que não foi possível um balanceamento total dos estímulos, já que alguns grupos não tiveram todas as pseudopalavras validadas<sup>4</sup>.

### 2.4 Quantidade de dados

O experimento contou, inicialmente, com 309 pseudopalavras<sup>5</sup>, que foram lidas por 34 falantes nativos do português brasileiro, totalizando 10.506 estímulos. Destes, foram excluídos 113 (1,1%) em decorrência inadequações nas produções<sup>6</sup>, resultando em 10.393 estímulos lidos.

### 2.5 Desenho experimental

Os estímulos foram exibidos aos participantes por meio de uma apresentação de *slides*, com uma palavra por tela. Todas as pseudopalavras foram escritas seguindo a ortografia da língua portuguesa, sem qualquer marcação de acento gráfico. É preciso destacar que é sabido que a forma escrita com ou sem gráfico dá pistas aos falantes em relação à sílaba proeminente, como em *sábia*, *sabia* e *sabiá*. Contudo, adotou-se tal metodologia, tendo em vista que diversos estudos prévios demonstram uma maior

---

<sup>4</sup> É preciso destacar que a decisão de excluir os estímulos não validados não se deu de forma aleatória, tendo em vista que as estatísticas que realizaram a análise inferencial deste trabalho optaram por excluir os dados não validados. Segundo elas, para que a hipótese de investigação fosse analisada, era preferível desbalancear os grupos experimentais do que incluir dados não validados. Os grupos mais afetados foram, dessa forma, os grupos similares, G1 e G3, resultando em 67 pseudopalavras no G1; 86 no G2; 65 no G3 e 91 no G4.

<sup>5</sup> Para mais detalhes da contabilização dos estímulos, consulte Benevides (2022).

<sup>6</sup> Os estímulos excluídos decorrem de troca de segmentos, (des)vozeamento, inserção de sílabas ou segmentos, mudança na ordem das sílabas etc.

neutralidade na leitura de palavras escritas do que na produção de palavras ouvidas ou formadas a partir de imagens (Benevides, 2017).

Para minimizar possíveis interferências do conhecimento que os falantes têm em relação ao sistema ortográfico da língua portuguesa, os participantes foram orientados de que as palavras expostas na tela do computador não apresentavam qualquer marcação de diacríticos (acentuação, cedilha etc.) em decorrência de falhas do programa no reconhecimento de caracteres. A tarefa experimental foi realizada *on-line*, com o compartilhamento da tela do computador da pesquisadora com os participantes. A chamada de vídeo e áudio foi realizada por meio de aplicativos como o *Zoom* e o *Google Meet*, com captação de áudio/vídeo realizada através do *notebook NP Compact*. Todos os participantes deste estudo autorizaram previamente a gravação de áudio/vídeo e o seu uso em meios acadêmicos.

## 2.6 Método de análise dos dados

Os dados foram submetidos a análises descritivas e inferenciais no *software R* (R Core Team, 2015), a partir da avaliação de um estatístico. O modelo utilizado na análise foi *modelo de regressão logística multinomial misto*, assumindo a paroxítona como categoria de referência.

## 3 RESULTADOS E ANÁLISE

Os resultados gerais da tarefa experimental compreenderam 4.569 (44%) de produções oxítonas, 5.436 (52,3%) de paroxítonas e 388 (3,7%) de proparoxítonas, conforme exposto na Tabela 1.

Tabela 1 - Frequência e proporção de tonicidade de produção nos dados gerais.

Tonicidade de Produção	n (%)
Oxítone	4.569 (44%)
Paroxítone	5.436 (52,3%)
Proparoxítone	388 (3,7%)
<b>Total</b>	<b>10.393 (100%)</b>

Fonte: Benevides (2022, p. 143).

Esses resultados gerais já trazem indícios de dois fatos relevantes: i. embora a paroxítona seja o acento predominante, as oxítonas também emergiram de forma significativa; e ii. as proparoxítonas não podem ser consideradas um padrão excepcional da língua, tendo em vista que, até mesmo em um estudo experimental, elas foram produzidas.

Para além dos resultados mais abrangentes, interessa-nos, nesta pesquisa, avaliar se a similaridade fonológica desempenha algum papel na atribuição do acento em português. Os dados quantitativos, apresentados na Tabela 2, ajudam-nos a observar que os grupos similares (G1 e G3) apresentaram taxas mais altas de acento oxítono e proparoxítono quando a palavra base também apresentava tal tonicidade, respectivamente. Em outros termos, 85% das pseudopalavras criadas a partir de uma

palavra real oxítone receberam também acento oxítone. Essa taxa mostra-se significativamente maior do que as pseudopalavras do grupo dissimilar, que só atingiram 51,1% ( $p < 0,001$ ). O efeito da similaridade parece ser ainda maior em relação às proparoxítonas, que atingiram uma taxa de 22,2% nos grupos similares em comparação a 4,4% nos grupos dissimilares ( $p < 0,001$ ). A análise inferencial seleciona a variável *similaridade fonológica* como uma variável relevante para a explicação do fenômeno em análise ( $p < 0,001$ ).

Tabela 2 - Proporção de tonicidade de produção por tonicidade da palavra base e por nível de (dis)similaridade.

Nível de Similaridade	Tonicidade de Produção = Tonicidade da Palavra Base		
	Oxítone	Paroxítone	Proparoxítone
<b>Similaridade (G1 e G3)</b>	85%	58,1%	22,2%
<b>Dissimilar (G2 e G4)</b>	51,1%	52,5%	4,4%

Fonte: Adaptação de Benevides (2022, 152).

É preciso destacar, ainda, que a análise inferencial indica que os grupos dissimilares (G2 e G4) favorecem o acento paroxítone. Este resultado mostra-se extremamente relevante, pois evidencia que a ausência de semelhança entre uma palavra real e uma pseudopalavra faz com que o falante atribua o padrão acentual *default* da língua - no caso, o paroxítone. Ou seja, uma possível interpretação é de que a ausência de um distrator, que direcione o acento para uma sílaba específica da palavra, como a similaridade fonológica, faz com que o indivíduo resgate o padrão acentual mais recorrente/robusto de sua língua, armazenado em seu léxico mental, e o atribua ao novo vocábulo. Há, assim, uma evidência importante de que o acento paroxítone é o padrão acentual do PB, já que é o mais frequente, o que leva à sua produtividade. Outra possível interpretação é de que o percentual em torno de 50% de acento oxítone e paroxítone nos grupos dissimilares pode dar indícios de uma atribuição aleatória, entretanto, como será evidenciado nos dados da Tabela 8, a sílaba final pode ser um fator explicativo.

Uma maneira de evidenciarmos como a similaridade fonológica desencadeia processos analógicos por meio da extensão do padrão acentual entre o vocábulo real e a pseudopalavra é detalharmos mais esses dados, a fim de analisarmos a relação entre atribuição acentual e tonicidade da palavra base, conforme demonstra a Tabela 3:

Tabela 3 - Frequência e proporção de tonicidade de produção por tonicidade da palavra alvo.

		Tonicidade de Produção			
		Oxítona	Paroxítona	Proparoxítona	Total
Tonicidade da Palavra Alvo	Oxítona	2.397 (63,6%)	1.331 (35,3%)	40 (1,1%)	<b>3.768</b> <b>(36,3%)</b>
	Paroxítona	1.863 (44%)	2.371 (56%)	65 (1,5%)	<b>4.299</b> <b>(41,4%)</b>
	Proparoxítona	309 (13,3%)	1.734 (74,5%)	283 (12,2%)	<b>2.326</b> <b>(22,4%)</b>
	Total	<b>4.569</b> <b>(44%)</b>	<b>5.436</b> <b>(52,3%)</b>	<b>388</b> <b>(3,7%)</b>	<b>10.393</b> <b>(100%)</b>

Fonte: Benevides (2022, p. 150).

Observe, como dito, que houve um maior percentual de atribuições de acento oxítono (63,6%) quando a pseudopalavra foi criada a partir de uma palavra com tonicidade oxítona. A probabilidade dela ser oxítona, entretanto, sofre uma queda de 75% e 90% quando a palavra base é paroxítona ou proparoxítona, respectivamente. Sob a mesma ótica, a probabilidade da palavra ser paroxítona aumenta se a palavra base for paroxítona ou proparoxítona ( $p < 0,001$ ). Há, ainda, 400% mais chances da pseudopalavra ser proparoxítona se ela tiver sido criada a partir de uma palavra proparoxítona.

Antes de adentrarmos na análise a respeito da frequência lexical da palavra base, cabe uma reflexão sobre a composição dos grupos experimentais: a fim de possibilitar que houvesse todos os contrastes de similaridade e frequência, criamos quatro grupos (G1 - Similar de alta frequência; G2 - Dissimilar de alta frequência; G3 - Similar de baixa frequência; e G4 - Dissimilar de baixa frequência). Contudo, ao observamos os resultados, expostos na Tabela 4, questionamo-nos se, de fato, tal divisão se mostra válida, tendo em vista que o pressuposto para a pseudopalavra ser enquadrada na categoria dissimilar é não apresentar semelhança com uma palavra real, de forma que não parece um raciocínio válido pensar que a frequência da palavra real que não se assemelha à pseudopalavra pode a afetar. Se tal raciocínio estiver correto, como julgamos que o seja, a melhor maneira de analisarmos o papel da frequência lexical da palavra base é isolar apenas o G1 dos demais grupos, como apresentado na Tabela 5:

Tabela 4 - Frequência e proporção de tonicidade de produção por tonicidade da palavra alvo e por nível de frequência.

Nível de Similaridade	Tonicidade da Palavra Alvo	Tonicidade de Produção			
		Oxítona	Paroxítona	Proparoxítona	Total
Grupos de Alta Frequência (1 e 2)	Oxítona	1.106 (62,1%)	667 (37,4%)	9 (0,5%)	<b>1.782</b> <b>(17,1%)</b>
	Paroxítona	950 (41,6%)	1.298 (56,8%)	38 (1,7%)	<b>2.286</b> <b>(22%)</b>
	Proparoxítona	53 (4,9%)	889 (82,5%)	136 (12,6%)	<b>1.078</b> <b>(10,4%)</b>
Grupos de Baixa Frequência (3 e 4)	Oxítona	1.291 (65%)	664 (33,4%)	31 (1,6%)	<b>1.986</b> <b>(19,1%)</b>
	Paroxítona	913 (45,4%)	1.073 (53,3%)	27 (1,3%)	<b>2.013</b> <b>(19,4%)</b>
	Proparoxítona	256 (20,5%)	845 (67,7%)	147 (11,8%)	<b>1.248</b> <b>(12%)</b>
	<b>Total</b>	<b>4.569</b> <b>(44%)</b>	<b>5.436</b> <b>(52,3%)</b>	<b>388</b> <b>(3,7%)</b>	<b>10.393</b> <b>(100%)</b>

Fonte: Benevides (2022, p. 153).

Tabela 5 - Proporção de tonicidade de produção por tonicidade da palavra base e por nível de frequência.

Nível de Frequência	Tonicidade de Produção = Tonicidade da Palavra Base		
	Oxítona	Paroxítona	Proparoxítona
<b>G1 - Similar de alta frequência</b>	91,6%	60,5%	21,5%
<b>Demais grupos (G2, G3 e G4)</b>	60,7%	53,4%	10,5%

Fonte: Benevides (2022, p. 206).

Observe, a partir da Tabela 5, que há significativamente mais atribuições oxítonas ( $p < 0,0001$ ) quando a palavra base é oxítona no grupo similar do que nos demais grupos - o mesmo vale para as demais tonicidades. Tal fato traz evidências de que a frequência da palavra base é um fator relevante na atribuição de tonicidade nesses vocábulos, podendo atuar, como pressuposto pelos Modelos baseados no Uso, como um mediador no acesso lexical. Isto é, a alta frequência das palavras faz com que elas tenham conexões mais robustas e, conseqüentemente, estejam mais disponíveis para serem acessadas pelos falantes, permitindo que os seus padrões sejam mais facilmente reproduzidos em novos vocábulos.

Vamos, agora, observar os resultados quanto à sílaba final. A Tabela 6 apresenta a frequência e a proporção de tonicidade de produção de acordo com a sílaba final da pseudopalavra, se CV ou CVC. Esta análise se faz necessária a fim de

avaliarmos se a forte correlação entre sílaba final e padrão acentual, herdada do latim, se mostra ainda um atrator importante da língua no que se refere à atribuição acentual.

Tabela 6 - Frequência e proporção de tonicidade de produção por sílaba final da pseudopalavra.

		Tonicidade de Produção			Total
		Oxítona	Paroxítona	Proparoxítona	
Sílaba Final da Palavra	CV-CV-CV	570 (10,6%)	4.518 (84,4%)	267 (5%)	5.355 (51,5%)
	CV-CV-CVC	3.999 (79,4%)	918 (18,2%)	121 (2,4%)	5.038 (48,5%)
	<b>Total</b>	<b>4.569 (44%)</b>	<b>5.436 (52,3%)</b>	<b>388 (3,7%)</b>	<b>10.393 (100%)</b>

Fonte: Benevides (2022, p. 151).

Note que há um maior número de atribuições de acento oxítono quando a pseudopalavra termina com CVC: 79,4% em contraste com 10,6% ( $p < 0,0001$ ). De modo semelhante, há mais atribuições paroxítonas quando a pseudopalavra termina com CV (84,4%). Para as proparoxítonas, o percentual é maior com CV (5%) em detrimento de CVC (2,4%). Este resultado é fundamental de ser comentado, visto que a literatura linguística, de modo geral, prediz que o acento proparoxítono não emerge quando a palavra termina em sílaba ramificada. Neste caso, boa parte deste resultado (2,4%) refere-se à emergência do acento proparoxítono em pseudopalavras que tinham como palavra base a palavra real *Júpiter*, explicitando, dessa forma, que a similaridade fonológica mostrou-se um fator mais relevante do que a baixa frequência de exemplares na língua com esse padrão (CVC final com tonicidade proparoxítona).

Observamos, até aqui, que as três variáveis isoladas são relevantes para a atribuição do acento em português brasileiro. Vamos, então, avaliar a interação entre elas. A Tabela 7 apresenta os dados referente à frequência e à proporção de tonicidade de produção por grupo experimental, o qual correlaciona a variável similaridade com frequência lexical.

Tabela 7 - Frequência e proporção de tonicidade de produção por grupo experimental.

Grupo Experimental	Tonicidade de Produção/Tonicidade da Palavra Alvo		
	Oxítona/ Oxítona	Paroxítona/ Paroxítona	Proparoxítona/ Proparoxítona
<b>Grupo 1</b> Similar de alta frequência	91,6%	60,5%	21,5%
<b>Grupo 2</b> Dissimilar de alta frequência	46,7%	53,2%	3,7%

<b>Grupo 3</b> <b>Similar de baixa frequência</b>	79,9%	55,2%	23%
<b>Grupo 4</b> <b>Dissimilar de baixa frequência</b>	55,4%	51,7%	4,9%

Fonte: Benevides (2022, p. 148).

Note que a correlação entre os fatores só reforça os resultados individuais de cada uma das variáveis: nos grupos similares (G1 e G3), há um maior percentual de atribuições oxítonas e proparoxítonas quando a palavra base é, respectivamente, oxítônica e proparoxítona; o percentual dos grupos dissimilares (G2 e G4) é muito semelhante, em torno de 50%, reforçando a nossa (re)análise de que esses grupos, na verdade, constituem o mesmo grupo, já que o pressuposto da formação desses grupos é de que os falantes não conseguem recuperar informações da palavra base. Ao unirmos esses grupos, temos 51,1% de produções oxítonas e 52,5% de proparoxítonas, percentuais que podem indicar a aleatoriedade. Esta afirmação só se mostra equivocada quando observamos os resultados dos grupos experimentais também por sílaba final, conforme exposto na Tabela 8.

Ao expandirmos o olhar, verificamos que, na verdade, essa atribuição, aparentemente aleatória, é motivada pela composição da sílaba final: quando ramificada, há um favorecimento do acento oxítono em todos os grupos e, quando não ramificada, há um favorecimento do acento proparoxítono também em todos os grupos. O percentual em torno de 50% pode decorrer, na verdade, de termos vocábulos com as duas estruturas de sílaba final, CV e CVC. O fator central desencadeador do acento nos grupos dissimilares, portanto, parece ser a sílaba final. Note, entretanto, que o percentual de atribuições de acento oxítono com CVC e proparoxítono com CV cai nos grupos similares: em média, 74,3% de oxítonas com CVC e 76,7% de proparoxítonas com CV, ao passo que os grupos dissimilares apresentam, em média, 84,5% de oxítonas com CVC e 88,6% de proparoxítonas com CV. A explicação para tal queda só pode ser o fator que distingue esses grupos, a similaridade fonológica. Há, para esses casos, mais de um fator que interage: similaridade fonológica, frequência lexical e sílaba final.

Tabela 8 - Frequência e proporção de tonicidade de produção por grupo experimental e por sílaba final.

Grupo Experimental	Sílabas Final da Palavra	Tonicidade de Produção		
		Oxítônica	Paroxítona	Proparoxítona
<b>G1 - Similar Frequente</b>	<b>CV-CV-CV</b>	45 (4,3%)	873 (83,4%)	129 (12,3%)
	<b>CV-CV-CVC</b>	950 (78,1%)	259 (21,3%)	7 (0,6%)
<b>G2 - Dissimilar Frequente</b>	<b>CV-CV-CV</b>	120 (7%)	1.549 (90,6%)	40 (2,3%)

	<b>CV-CV-CVC</b>	994 (84,7%)	173 (14,7%)	7 (0,6%)
<b>G3 - Similar de Baixa Frequência</b>	<b>CV-CV-CV</b>	228 (24,9%)	641 (70%)	47 (5,1%)
	<b>CV-CV-CVC</b>	897 (70,4%)	295 (23,1%)	83 (6,5%)
<b>G4 - Dissimilar de Baixa Frequência</b>	<b>CV-CV-CV</b>	177 (10,5%)	1.455 (86,5%)	51 (3%)
	<b>CV-CV-CVC</b>	1.158 (84,3%)	191 (13,9%)	24 (1,7%)

Fonte: Benevides (2022, p. 155).

É possível constatar, dessa maneira, que todos esses fatores contribuem, em alguma medida, para a atribuição do acento em pseudopalavras. Ao nosso ver, esses, possivelmente, não são os únicos. É preciso mais estudos para mapear e avaliar o papel que outras variáveis podem desempenhar em um fenômeno linguístico tão complexo como o acento lexical.

#### 4 CONSIDERAÇÕES FINAIS

Diante das inúmeras pesquisas que buscam determinar um único fator como motivador para a atribuição do acento, este estudo, de caráter experimental, tem como premissa a atuação de múltiplos fatores, à luz dos Modelos baseados no Uso. Os fatores investigados foram: similaridade fonológica, frequência de ocorrência do item lexical e sílaba final. Os resultados experimentais evidenciam a importância de todos eles na determinação do local em que o acento incide.

A relevância de mais de um fator só reforça a hipótese de análise desta pesquisa de que os padrões fonológicos, incluindo, neste caso, os padrões acentuais, são abstraídos e categorizados na gramática fonológica dos falantes. Por estarem disponíveis nas representações mentais (ou seja, no seu léxico mental), podem ser, e são, replicados em vocábulos novos da língua - neste caso, em pseudopalavras, a partir de processos analógicos. Nesta pesquisa, o mecanismo desencadeador desses processos foi a similaridade fonológica entre os vocábulos, sendo que a frequência de ocorrência do item lexical mostra-se um potencializador da analogia. Os dois fatores interagem com a sílaba final, que é um *template* extremamente relevante e ativo na língua, sendo o principal determinante, por exemplo, do local em que o acento incide quando não há similaridade fonológica ou outros fatores interagindo.

É possível constatar, dessa forma, que os falantes acessam informações segmentais e as consideram no momento de atribuir tonicidade a um vocábulo, demonstrando que os níveis segmentais e métricos estão relacionados nas representações mentais (não devendo ser concebidas de forma independente). Em outros termos, os falantes fazem processos analógicos segmentais e acentuais, o que indica que o acento está associado à palavra, isto é, é lexical.

Em síntese, os resultados desta pesquisa estão em consonância com a hipótese de Cantoni (2013) de que o acento primário em português brasileiro traz resquícios de

uma herança latina que associou sílaba final com padrão acentual, por isso, a sua maior influência em pseudopalavras dissimilares; e, principalmente, a defesa de que o acento é lexical, visto que ele está associado a um vocábulo que tem seu padrão abstraído e replicado em palavras fonologicamente semelhantes.

## REFERÊNCIAS

- Benevides AL. O acento primário em pseudopalavras: uma abordagem experimental [dissertação]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2017.
- Benevides AL. O acento em pseudopalavras [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2022.
- Benevides AL, Guide BF. Corpus ABG. *Texto Livre: Linguagem e Tecnologia*. 2017;10 (1):139-163.
- Bisol L. O acento e o pé métrico. *Letras de Hoje*. 1994;29(4):25-36.
- Bybee J. *Phonology and language use*. Cambridge: Cambridge University Press; 2001.
- Bybee J. From usage to grammar: the mind's response to repetition. *Scholarly journals online*; 2006. p. 711-732.
- Bybee J. *Language, usage and cognition*. Cambridge: Cambridge University Press; 2010.
- Câmara Jr. JM. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro: Simões; 1953.
- Câmara Jr. JM. *Estrutura da língua portuguesa*. 34.<sup>a</sup> ed. Petrópolis: Editora Vozes; 2001 [1970].
- Cantoni MM. O acento no português brasileiro: uma abordagem experimental [tese]. Belo Horizonte: Estudos Linguísticos da Faculdade de Letras, Universidade Federal de Minas Gerais; 2013.
- Coltheart M. Modeling reading: the dual-route approach. In: Snowling MJ, Hulme C, editores. *The science of reading: a handbook*. Oxford: Blackwell; 2005. p. 6-23.
- Coltheart M, Rastle K. Serial processing in reading aloud: evidence for dual-route models of reading. *Journal of Experimental Psychology*. 1994;20(6):1197-1211.
- Cristóvão-Silva T. Intervocalic palatal glides in Brazilian Portuguese. *Revista de Estudos Linguísticos*. 1995;2(4):5-20.
- Kemmer S, Barlow M. Introduction: a usage-based conception of language. In: Barlow M, Kemmer S. *Usage based models of language*. Califórnia: CSLI Publications; 1999. p. vii-xxviii.
- Langacker R. A dynamic usage-based model. In: Barlow M., Kemmer S. *Usage-based models of language*. Califórnia: CSLI Publications; 1999. p. 1-63.
- Langacker R. *Cognitive grammar: a basic introduction*. Nova Iorque: Oxford University Press; 2008.
- Langacker R. *Essentials of cognitive grammar*. Nova Iorque: Oxford University Press; 2013.
- Lee SH. *Morfologia e fonologia lexical do português do Brasil* [tese]. Campinas: Instituto de Estudos da Linguagem, Universidade Estadual de Campinas; 1995.
- Lee SH. O acento primário no português: uma análise unificada na teoria da otimalidade. In: Araújo GA, organizador. *O acento em português: abordagens fonológicas*. São Paulo: Parábola Editorial; 2007. p. 121-144.
- Magalhães JS. *O plano multidimensional do acento na teoria da otimalidade* [tese]. Porto Alegre: Faculdade de Letras, Pontifícia Universidade Católica do Rio Grande do Sul; 2004.
- Massini-Cagliari G. *Do poético ao linguístico no ritmo dos trovadores: três momentos da história do acento*. São Paulo: Cultura Acadêmica; 1999.
- Mateus MHM. O acento em palavra em português: uma nova proposta. *Boletim de Filologia*. 1983;27:211-229.

Pierrehumbert JB. Exemplar dynamics: word frequency, lenition and contrast. In: Bybee J, Hopper P, editores. Frequency effects and the emergence of linguistic structure. Amsterdam: John Benjamins; 2001. p. 137-157.

Protopapas A, et al. Lexical and default stress assignment in reading Greek. Journal of Research in Reading. 2006;29(4):418-432.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing [programa de computador]. Vienna, Austria. 2015. Disponível em: <http://www.R-project.org>.

Rastle K, Coltheart M. Lexical and nonlexical print-to-sound translation of disyllabic words and nonwords. Journal of Memory and Language. 2000;42:342-364.

Wetzels WL. Consoantes palatais como geminadas fonológicas no Português Brasileiro. Revista de Estudos Linguísticos. 2000;9(2):5-15.

Wetzels WL. Primary word stress in Brazilian Portuguese and the weight parameter. Journal of Portuguese Linguistics. 2007;5(2):9-58.

**APÊNDICE - Pseudopalavras por grupo experimental com suas respectivas frequências**

Grupo 1 - Pseudopalavras similares de alta frequência

CV-CV-CV			
Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Paraná	77.547	tarana, patana, parafa
			perana, parena, parane
	Canadá	25.505	fanada, cadada, canaga
			conada, canida, canadu
	Comitê	33.263	jomite, copite, comife
			camite, comote, comitu
Paroxítona	Cidade	525.295	zidade, cicade, cidafe
			cedade, cidide, cidadu
	Tabela	144.478	babela, tafela, tabeda
			tobela, tabala, tibela
	Figura	372.894	digura, fidura, figuba
			fagura, figora, figuri
Proparoxítona	Número	842.189	lumero, nupero, numeco
			nimero, numaro, namero
	Século	219.623	neculo, sevulo, secuno
			siculo, saculo, secule
Método	317.435	setodo, mepodo, metopo	

			mutodo, metedo, metode
--	--	--	------------------------

**CV-CV-CVC**

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Capital	197.974	lapital, cabital, capifal
			copital, caputal, capitol
	Diretor	140.646	siretor, difetor, diredor
			daretor, dirotor, diretur
	Natural	114.030	datural, nabural, natupal
			nitural, nataral, naturol
Paroxítona	Possível	300.554	bocivel, pobivel, pocibel
			picivel, pocevel, pocivol
	Caráter	115.284	darater, cacater, caraper
			curater, caruter, carator
	Difícil	104.458	zifícil, dipícil, difíbil
			dafícil, difucil, dificol

Grupo 2 - Pseudopalavras dissimilares de alta frequência

**CV-CV-CV**

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
Oxítona	Paraná	77.547	lapaca, dabaza, zavaba
			poruna, piruno, poranu
	Comitê	33.263	zofire, zoteme, dovife
			camota, camate, comato
	Canadá	25.505	fapama, lacafa, bavaga
			conedo, conado, cundo
Paroxítona	Medida	380.157	tefifa, lenipa, tepiva
			midode, madda, medoda
	Passado	211.473	narafo, nagafu, dalapo
			piceda, pucido, pacido

	<b>Futuro</b>	137.205	pucuto, nubuvo, dumufo futira, fetara, fitare
<b>Proparoxítona</b>	<b>Título</b>	183.459	sibumo, mivipo, lirufu titale, tetola, tatolo
	<b>Código</b>	210.472	nofilo, zofisu, bocito cadegu, coduga, cudoga
	<b>Década</b>	180.750	tefaza, sedafa, revafa dacodo, decode, ducoda

**CV-CV-CVC**

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
<b>Oxítona</b>	<b>Federal</b>	471.606	pevebal, befipal, mefebal faduril, fodoral, fidarel
	<b>Militar</b>	115.102	pitidar, dipinar, sipivar mulotar, malutor, molitur
	<b>Popular</b>	92.011	mogutar, lonutar, zoricar pipalor, papilor, pepalir
<b>Paroxítona</b>	<b>Possível</b>	300.554	gonitel, lomipel, rovinel pacevol, pecival, puceval
	<b>Caráter</b>	115.284	fanaler, zamafer, napaver corutar, carotir, coruter
	<b>Difícil</b>	104.458	mizidil, fidimil, fepibil dufical, dafocol, dafical

Grupo 3 - Pseudopalavras similares de baixa frequência

**CV-CV-CV**

Tonicidade	Palavra Alvo	Frequência	Pseudopalavra
<b>Oxítona</b>	<b>Paletó</b>	1.591	taletto, pabeto, palevo pileto, paluto, poletto
	<b>Javali</b>	690	bavali, japali, javari

			juvali, javoli, javala
	<b>Cabaré</b>	196	nabare, camare, cabate cobare, cabure, cabara
<b>Paroxítona</b>	<b>Safári</b>	932	nafari, savari, safapi sefari, safuri, safara
	<b>Pateta</b>	417	lateta, pafeta, pateda puteta, patuta, patetu
	<b>Sacola</b>	311	vacola, sanola, sacoda sucola, sacala, sacolo
<b>Proparoxítona</b>	<b>Pétala</b>	1.928	tetala, pefala, petafa patala, petula, petale
	<b>Cócega</b>	261	docega, cobega, coceca cacega, cucega, cocegu
	<b>Tísica</b>	245	disica, tirica, tisiba tasica, tisaca, tisoca

**CV-CV-CVC**

<b>Tonicidade</b>	<b>Palavra Alvo</b>	<b>Frequência</b>	<b>Pseudopalavra</b>
<b>Oxítona</b>	<b>Secador</b>	1.543	becador, segador, secapor secader, secodor, secadar
	<b>Chafariz</b>	606	nafariz, xabariz, xafadiz xofariz, xafuriz, xafaraz
	<b>Recital</b>	143	fecital, relital, recival rucital, racital, recitol
<b>Paroxítona</b>	<b>Cadáver</b>	191	ladaver, cazaver, cadaber codaver, cadiver, cadavar
	<b>Mutável</b>	933	zutavel, muvavel, mutazel mitavel, mutevel, mutaval
	<b>Lavável</b>	204	davavel, labavel, lavatel

			lavavil, lavevel, lavaval
<b>Proparoxítona</b>	<b>Júpiter</b>	1.904	tupiter, jutiter, jupiner
			japiter, jupiter, jupitor

Grupo 4 - Pseudopalavras dissimilares de baixa frequência

**CV-CV-CV**

<b>Tonicidade</b>	<b>Palavra Alvo</b>	<b>Frequência</b>	<b>Pseudopalavra</b>
<b>Oxítona</b>	<b>Cabaré</b>	1.020	lafape, barane, vatabe
			cabori, cobiro, cubere
	<b>Sucuri</b>	348	gumuli, govubi, vunuti
			socarú, sicori, sicora
	<b>Camelô</b>	184	bareno, navedo, lacepo
			cumile, comalo, camola
<b>Paroxítona</b>	<b>Mazela</b>	188	davena, fabeta, capeca
			mazole, muzola, mezilo
	<b>Farofa</b>	842	lamofa, dacuga, tagota
			firafe, farefo, firufo
	<b>Jaleco</b>	260	narebo, pamezo, tafeno
			juluca, jalica, jeluci
<b>Proparoxítona</b>	<b>Sádica</b>	250	gafila, tapina, lafipa
			soduca, sedaco, sedoca
	<b>Pêssego</b>	197	lemifo, demefo, zedeno
			pucega, pacigo, piciga
	<b>Súbito</b>	189	bufino, nufilo, zutino
			subata, sebeta, sobite

**CV-CV-CVC**

<b>Tonicidade</b>	<b>Palavra Alvo</b>	<b>Frequência</b>	<b>Pseudopalavra</b>
<b>Oxítona</b>	<b>Funeral</b>	1.975	mupeval, lufipal, puvebal
			fanirel, fenarel, finoral

	<b>Comedor</b>	232	podecor, votepor, motefor
			cumodor, cameder, motefor
	<b>Liminar</b>	193	sidigar, pifitar, ritecar
			lomenar, lemunor, lomenir
<b>Paroxítona</b>	<b>Cadáver</b>	191	lamater, nafapir, dapamer
			codivar, cudivor, cadovur
	<b>Mutável</b>	933	sulabel, ruvatel, vupadel
mitoval, mitavol, motevel			
	<b>Lavável</b>	204	salanel, fadapel, fapatil
			luvivol, livoval, levavol
<b>Proparoxítona</b>	<b>Júpiter</b>	1.904	luniver, tuzemer, gudiler
			japutar, jepatar, jopater



## Sândi vocálico externo em português brasileiro: elisão, ditongação e coalescência

*Brazilian Portuguese external sandhi:  
elision, diphthongization and coalescence*

Magnun Rochel Madruga\*

Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

Guilherme Gonçalves\*\*

Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

**Resumo:** O estudo apresenta resultados experimentais de uma análise fonético-acústica da elisão da vogal baixa /a/ por sândi vocálico externo no português brasileiro (PB) (BISOL, 1992). A pesquisa adotou o protocolo de leitura de sentenças-veículo, conforme descrito em Albano (2001) e Meneses (2016). Participaram do experimento oito falantes nativos do PB. O encontro vocálico foi controlado para manter o contexto favorecedor da elisão: vogal baixa /a/ átona em final de palavra seguida de [i] ou [u] átonos. A estrutura de juntura favorável à elisão resultou em três padrões de produção: elisão, ditongo decrescente e coalescência. O padrão coalescente é inédito sobre o sândi em PB, no qual emerge uma vogal [e] ou [o] das sequências /a##i/ ou /a##u/. Definimos como coalescência casos em que a vogal resultante do sândi apresentava padrão formântico de monotongos com características das vogais médias [e] ou [o]. Além de interpretar os resultados experimentais sobre o sândi vocálico, apresentamos uma análise fonológica que (i) incorpora a coalescência como subprocesso de elisão e (ii) elimina a degeminação como processo de sândi vocálico externo. A proposta revela um sistema de operações fonológicas mais econômicas e mais ajustadas ao que se observa à realidade fonética.

**Palavras-chave:** Sândi vocálico externo. Coalescência. Elisão. Degeminação. Ditongação.

**Abstract:** The study presents experimental results of a phonetic-acoustic analysis of the elision of the low vowel /a/ by external vowel sandhi in Brazilian Portuguese (BP) (BISOL, 1992). The research adopted the target sentence reading protocol described in Albano (2001) and Meneses (2016). Eight native BP speakers participated in the experiment. The [V1 ## V2] contact was controlled to inquiry the context favoring elision: unstressed low vowel /a/ at the end of a word followed by unstressed [i] or [u]. The word-boundary structure resulted in three production patterns: elision, decreasing diphthongs and coalescence. Coalescence is an innovative result regarding sandhi in BP, in which a vowel [e] or [o] emerges from the sequences /a##i/ or /a##u/. We defined as coalescence cases in which the vowel resulting from sandhi presented a formant pattern of monophthongs with characteristics of the mid vowels [e] or [o]. In addition to interpreting the experimental results on vowel sandhi, we present a phonological analysis that (i) incorporates coalescence as a subprocess of elision and (ii) eliminates degemination as an external vowel sandhi process. The proposal reveals a system

---

\* Professor Adjunto, Faculdade de Letras, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil; [magnun@ufmg.br](mailto:magnun@ufmg.br)

\*\* Mestrando em Direito, Bacharel Letras e Direito, Universidade Federal de Ouro Preto (UFOP), Ouro Preto, MG, Brasil; [guigefd@gmail.com](mailto:guigefd@gmail.com)

of phonological operations that is more economical and better adjusted to what is observed in phonetic productions.

**Keywords:** External sandhi. Coalescence. Elision. Degemination. Diphthogization.

## 1 INTRODUÇÃO

O sândi vocálico externo (SVE) é um processo em que ocorre ressilabação na fronteira entre duas palavras fonológicas, cujas vogais átonas final e inicial encontram-se no domínio da frase fonológica (cf. Herslund, 1986; Bisol, 1992, 1993, 1994, 2000; Abaurre, 1996; Tenani, 2002, 2004, 2006, entre outros). Os processos de sândi ou de juntura são largamente conhecidos na literatura da área, tendo sido registrados em análises do sânscrito por Pāṇini. Kiparsky (2022) informa que, mesmo antes de Pāṇini, havia textos do sânscrito védico em que se postulavam regras sobre o sândi, com o objetivo de que os textos divinos pudessem ser reproduzidos oralmente em sua fidelidade com a forma oral da língua.

Na aprendizagem de uma segunda língua ou língua adicional, tais fenômenos são, inclusive, alvo de abordagens explícitas de ensino e configuram tópico importante para sistemas computacionais de conversão texto-fala, como é o caso da *connected speech* na aprendizagem de inglês e da *liaison* em francês (cf. Alameen Levis, 2015; Liakin et al., 2017).

A abordagem explorada neste texto não é pedagógica e tampouco se preocupa com uma implementação computacional; trata-se de uma abordagem teórico-descritiva. Para o português, afirma-se normalmente que o SVE é caracterizado pelo encontro de dois segmentos vocálicos em fronteira de palavras fonológicas sob o mesmo domínio prosódico e que pode ser alvo de três diferentes processos fonológicos, conforme proposto, de forma pioneira, por Bisol (1992), a saber: elisão, degeminação e ditongação. Vemos em (1) exemplos adaptados de Bisol (1992, 1993, 2000) e de Tenani (2006) para SVE em português brasileiro.

(1)

### a. Degeminação

menin[a ## a]legre	→	menin[a]legre
leq[i ## i]scuro	→	leq[i]scuro
vej[u ## u]sinas	→	vej[u]sinas

### b. Ditongação

pêsseg[u ## a]marelo	→	pêsseg[wa]marelo
camis[a ## u]sada	→	camis[aw]sada
contav[a ## i]stórias	→	contav[aj]stórias

### c. Elisão

merend[a ## i]scolar	→	merend[i]scolar
er[a ## u]sado	→	er[u]sado

No que diz respeito ao domínio prosódico de SVE, desde Bisol (1992, 1993), considera-se que sua ocorrência dá-se sempre no escopo de uma mesma frase

fonológica (cf. Tenani, 2006; Vianna, 2009; Paulino, 2016; Nogueira, 2007)<sup>1</sup>. Dessa forma, ainda que Abaurre (1996) e Tenani (2004) atestem que SVE também pode ser sensível ao acento frasal, capaz de bloquear sua ocorrência pela relação com informações rítmicas, entoacionais e sintáticas, ambas admitem que o domínio de SVE é a frase fonológica. Considerando que o encontro de vogais átonas ( $V_1 \## V_2$ ) é sempre o contexto de SVE, o seu bloqueio apresenta a seguinte relação com a estrutura acentual das vogais envolvidas, conforme exemplos adaptados de Bisol (1993, 19:22):

(2)

**a. Vogal Átona + Vogal Acentuada**

$V_1 \## V_2 \rightarrow$  bloqueio de DEG e EL

- |        |           |             |
|--------|-----------|-------------|
| 1. DEG | casa alta | *cas[zaw]ta |
| 2. EL  | fala isto | *fa[lis]to  |

**b. Vogal Acentuada + Vogal Átona**

$V_1 \## V_2 \rightarrow$  bloqueio de EL, mas não de DEG

- |        |               |                |
|--------|---------------|----------------|
| 3. DEG | vi estrelas   | ?[vis'trelɐs]  |
| 4. DEG | caju usado    | ?[kaʒu'zadõ]   |
| 5. EL  | babá elegante | *ba[be]legante |

**c. Vogal Acentuada + Vogal Acentuada**

$V_1 \## V_2 \rightarrow$  bloqueio de DEG, EL

- |        |         |          |
|--------|---------|----------|
| 6. DEG | vi isso | *[visu]  |
| 7. EL  | vi arte | *[var]te |

Para esta autora, nos casos de (2b-3,4), ocorrências de vogais idênticas sob condições de acento em  $V_1$ , pode ocorrer a degeminação. No entanto, usamos a notação (?) para indicar que essas produções parecem ser diferentes dos casos em que ocorre, de fato, a elisão, sendo, portanto, produções suspeitas. Nos casos (2c-6,7), na fala espontânea, parecer haver a produção de duas vogais, indicando a existência de vogais fonéticas, i.e. [vi.'i.su] e [ka.ʒu.u.'za.dõ] e não a silabificação exemplificada pela autora. Tal fato indicaria, portanto, que o acento na primeira vogal da juntura bloquearia também casos de DEG, e não apenas os de EL. Conforme a autora, mesmo havendo bloqueios, para todos os casos de acento em  $V_1$  ou  $V_2$ , a ditongação é sempre possível, seja ela crescente ou decrescente, embora a crescente seja a mais observada. De todo modo, vale notar aqui que os bloqueios evidenciam restrições condicionadas ao acento de umas das vogais para a ressilabação, que, segundo Bisol (1992), é o mecanismo controlador de sândi externo em português.

Assumindo a premissa de ressilabação como mecanismo operador de SVE e as restrições acentuais impostas para a sua ocorrência, como propõe Bisol (1993), este trabalho pretende (i) investigar experimentalmente o processo de sândi vocálico externo em português brasileiro, descrevendo acusticamente as vogais resultantes de SVE; e (ii) propor uma análise fonológica para o funcionamento do processo de SVE para o PB. O estudo de SVE, em uma abordagem experimental, justifica-se pela necessidade de uma descrição fonética das vogais resultantes em SVE de diferentes combinações fonotáticas de fronteira (e.g. Meneses, 2016, analisa apenas vogais

<sup>1</sup> Os processos de juntura que envolvem clíticos não estão no escopo deste trabalho, como por exemplo <do amigo> d[wa]amigo, embora os resultados de SVE sejam aparentemente os mesmos.

resultantes da sequência /a##i/. Além disso, uma análise baseada nos resultados experimentais permitirá uma compreensão mais precisa do funcionamento fonológico do fenômeno, permitindo o aprimoramento de interpretações disponíveis na literatura.

## 2 EXPERIMENTO

### 2.1 Hipóteses

De modo a responder os objetivos propostos, as hipóteses de trabalho são: (i) H1: a taxa de elocução não está associada ao sândi vocálico externo; ii) H2: a qualidade acústica das vogais em SVE é idêntica à das vogais de sílabas pretônicas; iii) H3: o *parsing* fonológico da vogal resultante de sândi é a pauta pretônica. Em relação à primeira hipótese, adicionou-se ao experimento três repetições de cada sentença, uma vez que, ao passo que o falante repete os estímulos, tende a aumentar a sua taxa de elocução. Albano (1999) e Meneses (2016), ao tratarem da elisão de /a/ em sândi, afirmam que o fenômeno é gradiente, isto é, que resultados fonéticos estão em função do tempo e das condições de produção, tal como a taxa de elocução e fronteiras prosódicas. Desse aspecto, decorre que, se as vogais tendem a se apresentar com qualidades foneticamente distintas, faz-se necessária descrição de suas características acústicas, mas, ao estarem parseadas como vogais pretônicas, devem apresentar características de tais (H2). A hipótese H3 segue Bisol (1996), para quem a sílaba resultante do sândi prosodiza-se como vogal pretônica. Por fim, para testar essa hipótese, apresentamos o desenho experimental nas seções que seguem.

### 2.2 Materiais e Método

Neste trabalho<sup>2</sup>, investigamos as características acústicas da vogal átona resultante do processo de elisão por SVE em sequências do tipo /V<sub>1</sub>##V<sub>2</sub>/ átonas. Realizou-se uma tarefa de leitura de sentenças nas quais se analisaram os parâmetros fonético-acústicos F1, F2 e duração. O objetivo do experimento é verificar as características acústicas da vogal resultante de sândi.

As gravações foram feitas em sala com isolamento acústico da Faculdade de Letras da Universidade Federal de Minas Gerais (FALE/UFMG), utilizando-se microfone do computador Macbook Air, 2019, 13 polegadas, modelo de duas saídas *thunderbolt 4*. Para segmentação dos dados e análise acústica, utilizou-se o programa de computador PRAAT, versão 6.2.22, (BOERSMA & WEENINCK, 2022). Os arquivos de áudio foram gravados em canal mono com frequência de amostragem de 44.100 Hz. As gravações foram salvas em formato de arquivo do tipo .WAV.

### 2.3 Participantes

Participaram do experimento 8 falantes nativos do PB naturais de Belo Horizonte, 4 do sexo feminino e 4 do sexo masculino, todos com idade entre 20 e 33. Nenhum deles reportou ter algum prejuízo auditivo ou de fala.

---

<sup>2</sup> A execução desta pesquisa conta com aprovação do Comitê de Ética em Pesquisa da UFMG, sob o número CAAE: 15189019.0.0000.5149.

## 2.4 Estímulos

Os estímulos do experimento foram construídos seguindo a estrutura de um sintagma nominal com sequência /V<sub>1</sub>##V<sub>2</sub>/ controlada conforme descrição estrutural da elisão definida por Bisol (1992). Sendo assim, para a escolha do sintagma, fixou-se V<sub>1</sub> com a vogal final átona /a/ e V<sub>2</sub> com as vogais /i/ ou /u/.

As palavras-alvo foram inseridas em sentenças-veículo do tipo “*Isso é conhecido como uma sequência-alvo por eles*”. Como estímulo controle, foram criados logatomas em que as vogais /i/ e /u/ pretônicas coincidiram com a sílaba-alvo do contexto de sândi, criando-se pares mínimos ortográficos, tal como utilizado em Albano (2001) e Meneses (2016). Logo, formaram-se pares como “*cabeça idônea*” e “*cabecidônea*”, em que o primeiro testa a elisão e o segundo é o estímulo controle. Considerando as características acústicas dos sons, para melhor acurácia na segmentação, procurou-se utilizar apenas consoantes oclusivas no *onset* da sílaba tônica na construção dos logatomas, e.g. *tabela##icônica vs. tabelicônica*, aos moldes de pares “*cara + idoso*” vs. “*caridoso*” e presentes em Albano (2001).

Na sílaba tônica da primeira palavra da sequência-alvo, foram selecionadas as sete vogais tônicas do PB /i, e, ε, a, ɔ, o, u/ e as cinco pretônicas foram inseridas no logatoma controle: /i, e, a, o, u/. As sequências /ti/ e /di/ foram evitadas em função do processo de palatalização das oclusivas no PB, cujas características acústicas podem afetar a produção plena da vogal (Meneses, 2016). Na Tabela 1, apresentamos os estímulos utilizados.

Tabela 1 - Estímulos utilizados no experimento.

Tipo de V2	Sintagma Alvo	Logatoma
/i/	pomposa ## ideia	pomposideia
	cultura ## idônea	culturidônea
	cabeça ## idônea	cabecidônea
	cantora ## icônica	cantor <sup>i</sup> cônica
	tabela ## icônica	tabelicônica
	maloca ## icônica	maloquicônica
	profana ## ideia	profanideia
/u/	nociva ## usina	nocivuzina
	postura ## usada	posturuzada
	vereda ## usada	vereduzada
	garota ## usada	garotuzada
	tabela ## usada	tabeluzada
	danosa ## usina	danozuzina
	catara ## usada	cataruzada

Fonte: Elaborado pelos autores.

O encontro vocálico /V<sub>1</sub>##V<sub>2</sub>/ foi controlado para manter-se a descrição estrutural da elisão: duas vogais átonas seguidas (e.g. [a ## V átona alta]). Logo, para cada combinação /a##i/ e /a##u/, foram criados sete pares diferentes (sintagma *v.* logatoma), uma vez que se controlou a qualidade da vogal tônica precedente à V<sub>1</sub> para fins de evitar-se uma fonotaxe igual entre a tônica e /a/ final (ver Tabela 1). Portanto, tendo em conta as restrições fonotáticas impostas, os estímulos totalizaram 84 palavras e sintagmas ([7 sintagmas-alvo + 7 logatomas-controle] x 2 tipos de V<sub>2</sub> x 3 repetições). O experimento contou ainda com aproximadamente 30% de palavras distratoras, totalizando 93 sentenças. A estrutura fonológica da sentença escolhida para os distratores considerou V<sub>1</sub> átona e V<sub>2</sub> tônica, tal como <cara##ávido>, estrutura descrita por Bisol (1993) como bloqueadora de EL e DEG (veja exemplo 2a). Além disso, as vogais V<sub>1</sub> e V<sub>2</sub> das palavras distratoras foram fixadas como vogais idênticas, diferenciando-se apenas pelo estatuto acentual do contato da juntura, sendo V<sub>2</sub> sempre tônica. A estrutura das palavras distratoras foi assim selecionada para que mantivéssemos o mesmo padrão de sentença, de modo a evitar mudanças prosódicas bruscas entre as sentenças-alvo e as distratoras.

## 2.5 Procedimento e Análise

A coleta dos dados foi feita em uma sala silenciosa, e a anotação de dados acústicos foi realizada no PRAAT (a partir de três níveis de segmentação: i) no primeiro nível, segmentaram-se a palavra controle e a sequência-alvo, por exemplo, <tabelusada> e <tabela#usada>; ii) na segunda camada de anotação, segmentaram-se a sílaba pretônica imediata à tônica da palavra controle e a sílaba resultante de SVE, *tabe[lu]sada*, de <tabela##usada>; iii) a terceira camada foi dedicada à segmentação das sílabas cujo resultado de SVE foi considerado um hiato ou um ditongo. No contexto fonológico V<sub>1</sub>##V<sub>2</sub> átonas, verificou-se também realização de hiatos, pausas e glotalizações, os quais somaram 24 casos na amostra de 328 produções (aprox. 7,32%). Pelo fato de que esses fenômenos não configuram processo de sândi, tais dados foram excluídos.

## 2.6 Análise Acústica

Para a análise acústica, extraíram-se os primeiros dois formantes das vogais (F1 e F2) resultantes de sândi no sintagma alvo e da vogal controle na palavra, bem como a duração da sílaba e da vogal. A extração dos parâmetros duração, F1 e F2 das vogais foi feita automaticamente, com revisão manual posterior, por *script*, que considera, para a extração de F1 e F2, o ponto médio da vogal, i.e., a porção estacionária dos formantes, enquanto para duração, o ponto inicial e final da anotação manual feita no PRAAT.

Com o objetivo de se obter uma análise formântica mais precisa, na análise realizada, o *script* utilizado estipula um teto de frequência de análise considerando o sexo do falante. Para falantes homens, que têm um trato articulatório maior, a frequência máxima de análise é 5000 Hz, enquanto, para falantes femininas, o teto de frequência é de 5500 Hz. Embora se utilize, neste estudo, somente os primeiros dois formantes, o *script* procura pelos três primeiros formantes, o que implica procurar frequências na faixa de F5 e F6 para assegurar a acurácia de F1 e F2 obtidos, considerando-se o tamanho do trato vocal como  $5c/(2 * \text{teto formântico})$ , em que *c* é a constante de velocidade do som (cf. Escudero et al. 2009 para detalhes).

### 3 DESCRIÇÃO E ANÁLISE DOS RESULTADOS

#### 3.1 Elisão, Coalescência e Ditongação

Os resultados observados para o processo de sândi vocálico externo são elisão, ditongação e coalescência. A elisão é entendida, neste trabalho, como o apagamento total do primeiro elemento da sequência de vogais, isto é, o apagamento de V<sub>1</sub>. A ditongação é a conversão da sequência de vogais de sílabas distintas, sendo uma delas é alta, em uma semivogal no domínio de apenas uma sílaba. A coalescência é uma fusão assimilatória redutiva, que funde dois elementos fonológicos e o segmento resultante expressa características intermediárias.

A elisão é o caso mais frequente, ocorrendo em torno de 50% nas três repetições. Não se observou um aumento do processo de elisão de /a/ conforme a repetição, mas uma estabilidade que varia entre 50-55% dos dados. Considera-se elisão os casos de apagamento da V<sub>1</sub> [a] e a permanência somente da V<sub>2</sub> [i, u], conforme inspeção espectrográfica e análise formântica, tal como <cantora###icônica> → cantor[i]cônica. A proporção dos padrões de sândi está disposta na Tabela 2.

Tabela 2 - Frequência absoluta e relativa dos resultados de SVE para cada repetição.

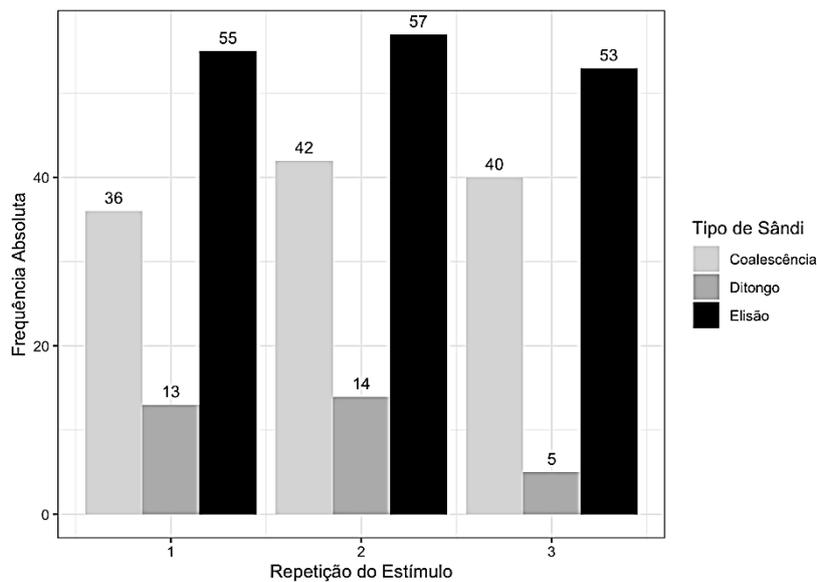
	Elisão	Ditongação	Coalescência	<b>Total</b>
Produção 1	55 (52%)	13 (12.4%)	36 (34.6%)	104 (100%)
Produção 2	57 (50.4%)	14 (12.4%)	42 (37.2%)	113 (100%)
Produção 3	53 (54.1%)	5 (5.1%)	40 (40.8%)	98 (100%)
<b>Total</b>	165 (52,4%)	32 (10.2%)	118 (37.4%)	315 (100%)

Fonte: Elaborado pelos autores.

A ditongação como processo resultante de SVE, por outro lado, tende a ser reduzida nas repetições um e dois. As produções ditongadas vão de 13% na primeira produção para 12% e 5% na segunda e terceira repetição. Os ditongos como resultantes de SVE corresponderam a apenas 10% do total. Em termos de realização fonética, observaram-se ditongos em que a vogal /a/ apresentou um alçamento, sendo realizada como [ə] seguidas dos glides [j, w] cuja origem são as vogais em V<sub>2</sub>, como em <tabela###icônica> → tabel[əj]cônica. Sobre a escolha do símbolo para essa vogal, optou-se por [ə] em função do seu valor médio de F1 estar na faixa de 500-560 Hz.

O resultado inovador deste trabalho é a caracterização da coalescência como um processo resultante de SVE. Tal achado é inédito na literatura da área sobre o PB, que tem sido descrito como uma língua cujos fenômenos de juntura restringem-se à elisão de V<sub>1</sub>, ditongação e degeminação (Bisol, 1992, 2000, 2002, Tenani, 2006, Vianna, 2009). O processo de coalescência caracteriza-se pela presença das vogais médias [e] e [o] como resultado de SVE. Essas vogais médias altas foram caracterizadas a partir dos seus valores médios de F1, os quais resultaram na faixa de 440-454 Hz para [e] e de 433-463 Hz para vogal posterior [o]. Nitidamente, a presença dessas vogais só pode ser definida pela fusão das vogais /a/ e /i, u/, uma vez que o resultado fonético expressa a conjunção da altura de /a/ com a posterioridade das vogais em V<sub>2</sub>.

Os resultados para a coalescência corresponderam a 37,4% do total e sua produção sofreu um leve aumento da primeira produção para a terceira, correspondendo a valores próximos a 35% na primeira repetição, 37% na segunda e a 41% dos casos na terceira. Casos de coalescência são aqueles que do encontro da V<sub>1</sub> /a/ e da V<sub>2</sub> /i, u/ resulta [e, o], como em: tabe[lo]sada <tabela##usada> e profa[ne]deia <profana##ideia>. Além disso, os ditongos encontrados tiveram sua vogal nuclear alçada para [ə], [e] ou [o] nos 32 casos observados (10,2% do total), indicando uma relação com a coalescência. Exemplos de ditongação crescente são: postur[ow]sada <postura##usada> e cabeç[ej]dônea <cabeça##idônea>.



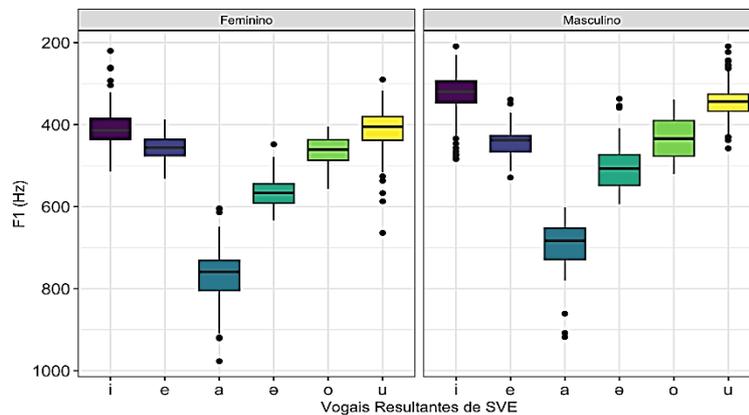
Fonte: Elaborado pelos autores.

Figura 3 - Frequência relativa dos processos de sândi externo.

Os resultados do teste de independência qui-quadrado para a associação entre número de repetições e tipo de resultado de sândi não revelaram significância estatística ( $\chi^2 = 4,24$ ,  $gl = 4$ ,  $p = 0,37$ ), sugerindo que a preferência de EL como fenômeno principal de SVE não parece estar associada, por exemplo, a um fator como taxa de elocução da fala. Se houvesse uma preferência de EL, DIT ou COAL em determinada repetição, ou mesmo uma tendência de aumento, observaríamos que EL tenderia a aumentar, mas o resultado obtido é a preservação da taxa de ocorrência dos fenômenos, com leves acréscimos para EL e COAL e um decréscimo não significativo para DIT.

Para a investigação da qualidade fonético-acústica das vogais de SVE, analisou-se principalmente o parâmetro F1 das vogais, cujos valores detalhados estão dispostos na Tabela 3. Isso foi necessário para caracterizar as vogais que eram percebidas, durante a análise manual, como vogais médias [e, o, ə]. Um modelo linear de efeitos mistos ajustado considerou F1 como variável resposta, GÊNERO e VOGAL RESULTANTE DE SÂNDI como covariáveis de efeitos fixos, enquanto as variáveis PARTICIPANTE e REPETIÇÃO foram definidas no modelo como variáveis de efeitos aleatórios. As suposições de normalidade dos erros e de homogeneidade das variâncias foram avaliadas graficamente através de gráficos Q-Q para probabilidade normal dos resíduos e pelo teste de Shapiro-Wilk, enquanto a homogeneidade das variâncias foi avaliada graficamente pelo gráfico de resíduos *versus* preditos e pelo teste

de Levene. O modelo geral, avaliado pela análise de *deviance* (Tipo II, teste de Wald), retornou, conforme esperado, que as covariáveis GÊNERO e VOGAL RESULTANTE DE SÂNDI são significativas, ao nível de  $\alpha = 5\%$ , para explicar F1 (VOGAL RESULTANTE DE SÂNDI:  $\chi^2 = 4070,04$ , gl = 5,  $p < 0,001$ ; GÊNERO:  $\chi^2 = 6,39$ , gl = 5,  $p = 0,011$ ).



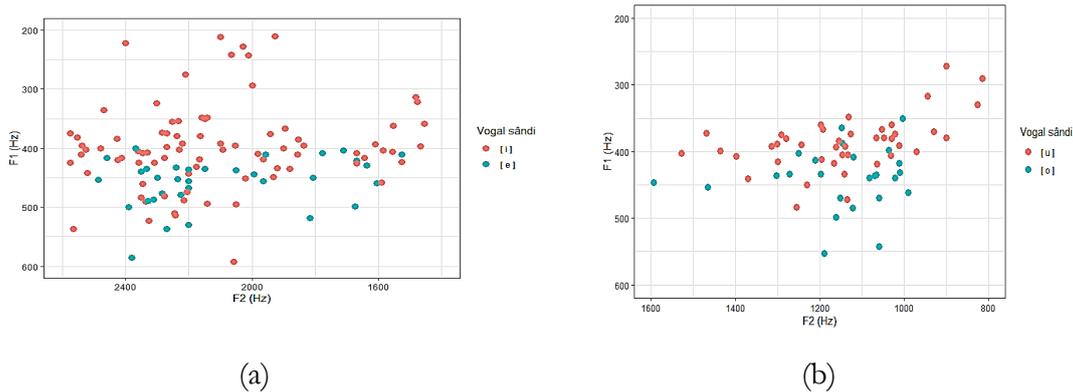
Fonte: Elaborado pelos autores.

Figura 4 - Vogais resultantes de SVE nos casos de EL, DIT e COAL agrupados por gênero.

Os valores de F1 confirmam a existência das vogais médias como [e, o, ə] no espaço acústico das vogais resultantes de SVE, tanto na realização ditongada quanto como monotongos. Caso os dados corroborassem completamente os processos propostos por Bisol (1992, 1996), esperava-se encontrar elisão de  $V_1$  ou uma ditongação do tipo [aw], mas o que se encontrou experimentalmente foi um alçamento de /a/ para [ə] ou mesmo [e] ou [o] nas ditongações, formando padrões como [əj, əw] ou [ej, ew, oj, ow].

Vale destacar, neste momento, a descrição da vogal [ə], *schwa*. Para ambos os sexos, a vogal média central ou *schwa* apresentou valores médios de F1 e F2 próximos, ficando em torno de 520-567 Hz para F1 e de 1504-1564 Hz para F2, e a escolha desse símbolo parece mais acurada do que o usualmente utilizado para o PB átona [ɐ] porque *schwas*, em geral, é resultado de processo de reduções fonológicas ou fonéticas. O intervalo de confiança de 95% para média da diferença entre [a - ə] é significativo ( $\beta = -1,57$ ;  $IC^{100(1-\infty)\%} = -1,77; -1,37$ ). O valor médio de F1 para [ə] foi de F1 = 519 Hz,  $s = 63$  Hz para mulheres, enquanto, para homens, o resultado foi F1 = 567 Hz,  $s = 40$  Hz. Como vemos, esse valor de F1 é muito próximo aos valores de uma vogal neutra do tipo *schwa* (F1 = 500 Hz; F2 = 1500 Hz) nas línguas do mundo, conforme previsto por Fant (1960). No catalão, por exemplo, que possui *schwa* em posição átona resultante da neutralização de /e, ɛ, a/, a literatura reporta valores de F1 até mesmo em faixa próxima a 600 Hz, enquanto, para o *schwa* átono final do inglês F1, de 665 Hz (cf. Fleming, 2009, p. 3, para ambas línguas). Embora não se pretenda fazer uma discussão sobre as diferenças entre os tipos de vogais centrais no português brasileiro (e.g. [ə] vs. [ɐ]), é importante descrever e reportar essa qualidade de vogal distinta do que se conhece sobre as realizações de /a/ nas condições de tonicidade e atonicidade.

O apagamento de  $V_1$  também ocorre, mas com consequências acústicas nas vogais resultantes, em que emerge [e, o, ə] como resultado da coalescência de altura e posterioridade de  $V_2$ . Conforme a Figura 4, os valores de F1 estão próximos 450 Hz para [e, o] tanto na produção de falantes mulheres quanto dos homens, indicando, portanto, que não se trata de uma elisão de /a/ em favor do [i] seguinte (cujos valores de F1 ficam em torno de 300 Hz), mas de um abaixamento que só pode ser explicado pela coalescência de /a/ com a vogal alta seguinte, seja ela [i] ou [u].



Fonte: Elaborado pelos autores.

Figura 5 - F1 x F2 das vogais [i, e] à esquerda (a) e de [u, o] à direita (b).

Nos painéis (a) e (b) da Figura 7, observam-se que as vogais médias são produzidas na porção inferior do gráfico (faixas de 400-600 Hz), enquanto as vogais altas, na porção superior, na faixa de 200 Hz a 400 Hz. Portanto, parece não haver dúvidas de que o processo de SVE envolve também um processo de coalescências de vogais. O teste de Tukey confirma que há uma diferença significativa entre os valores médios de F1 para [i-e] ( $\beta = -0,681$ ; EP = 0,056;  $z = -12,151$ ;  $p < 0,001$ ) e para [u-o] ( $\beta = -0,534$ ; EP = 0,049;  $z = -10,862$ ;  $p < 0,001$ ). Ao se inspecionar mais detalhadamente a vogal nuclear dos ditongos resultantes de SVE, encontrou-se um alçamento da vogal /a/, que caracterizamos como [ə], conforme Tabela 3. Como se verificou a existência de seis vogais fonéticas, o espaço acústico resultante de SVE pode ser descrito em forma de um Y, uma vez que a vogal [ə] tende a ocupar a porção central do espaço.

### 3.2 Espaço acústico para as vogais em SVE: a forma Y

Na Figura 6, podemos verificar que os espaços acústicos das vogais produzidas por homens e mulheres assumiram posições bem próximas ao reportado na literatura para as vogais pretônicas. Os valores observados para as médias de F1 e F2 são próximos ao que descrevem Barbosa et al. (2019) e Madruga et al. (2020) para as pretônicas, bem como muito semelhantes também a valores observados para essas mesmas vogais em posição tônicas, tal como reporta Escudero et al. (2009). A Tabela 3 sumariza os resultados deste estudo para as vogais em SVE.

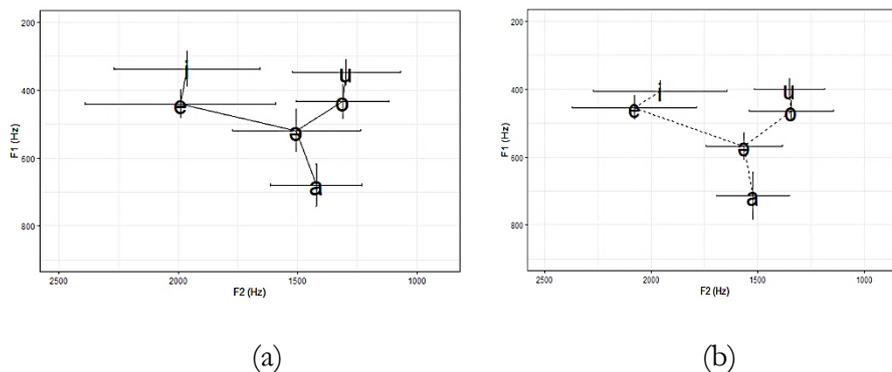
Tabela 3 - Média de F1 e F2 (em Hz) das vogais resultantes de SVE para cada gênero.

Gênero		[i]	[e]	[a]	[ə]	[o]	[u]
Feminino	F1	337 (42)	440 (52)	679 (63)	519 (63)	433 (51)	347 (38)
	F2	1962 (305)	1991 (398)	1422 (192)	1504 (267)	1311 (194)	1295 (228)
Masculino	F1	405 (31)	454 (34)	714 (70)	567 (40)	464 (32)	400 (32)
	F2	1958 (313)	2078 (291)	1521 (170)	1564 (178)	1342 (196)	1352 (164)

Nota: os valores entre parênteses referem-se aos desvios-padrões para F1 e F2.

Fonte: Elaborado pelos autores.

Na Figura 6 a seguir, são dispostas as seis vogais em seus pontos médios de F1 e F2, bem como seus respectivos desvios-padrões em forma de linhas horizontal e vertical.



Fonte: Elaborado pelos autores.

Figura 6 - Média e desvio-padrão do F1 e F2 das vogais resultado de sândi dos homens em (a) e mulheres em (b).

Para os valores de F2 das produções de [e] e [o] de ambos sexos, percebe-se que os pares de vogal [i, e] e [u, o] apresentam valores muito semelhantes para esse parâmetro. No entanto, Madruga et al. (2020) reportam F2 de [i] mais alto, tanto para homens quanto para mulheres, em torno de 2300 Hz. Para a vogal posterior [u], os autores atribuem F2 em torno de 1000 Hz, enquanto observamos vogais [u] resultantes de SVE com F2 na região dos 1300 Hz. Tal resultado aponta para o fato de que SVE parece exercer uma centralização do espaço acústico como um todo, uma vez que ambas as vogais altas são produzidas mais centralizadas, o que poderia funcionar como uma pista secundária para o processo de sândi externo. Obviamente, tal afirmação é uma hipótese que mereceria maior detalhamento, que, por questões de espaço, não será abordada no presente trabalho.

### 3.3 Duração das vogais em SVE

Em relação ao estatuto prosódico da sílaba resultante de sândi, testou-se a hipótese de que uma das possibilidades de *parsing* prosódico seria a sílaba realizar-se como uma sílaba postônica, uma vez que a vogal apagada está em posição átona final. Verificou-se que não há diferenças significativas entre a vogal do sândi e a vogal controle, ou seja, a vogal pretônica do logatoma. As estatísticas descritivas indicam uma duração média de 52 ms para a vogal [i] do sândi e de 54 ms para a vogal [i] controle. Esse resultado é um indicativo, portanto, de que a sílaba resultante do sândi prosodiza-se como vogal pretônica, tal como previsto por Bisol (1996). A Figura 7 demonstra que as durações apresentam uma distribuição muito próxima para cada vogal, seja [i] ou [u] em cada condição experimental.

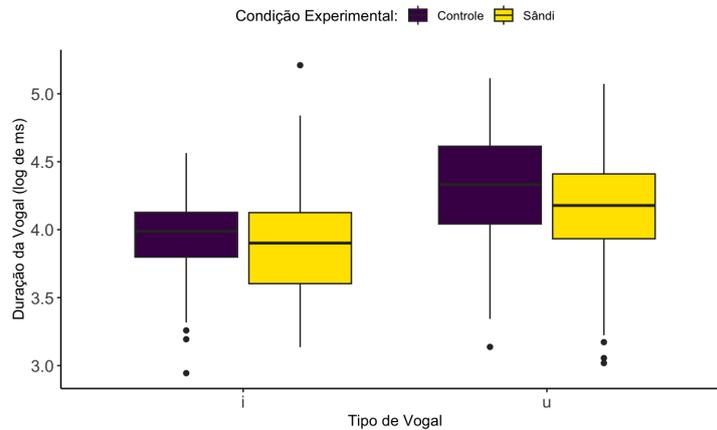
Tabela 4 - Média da duração (em milissegundos) das vogais resultado de sândi e controle e seus respectivos desvios-padrões.

	[i]	[u]
Sândi	52 (21)	68 (26)
Controle	54 (13)	80 (30)

Fonte: Elaborado pelos autores.

O modelo linear de efeitos mistos ajustado considerou a variável DURAÇÃO DA VOGAL em escala logarítmica como variável resposta e CONDIÇÃO EXPERIMENTAL (Controle ou Sândi) em interação com TIPO DE VOGAL como covariáveis de efeitos fixos, enquanto PARTICIPANTE e REPETIÇÃO foram definidas como variáveis de efeitos aleatórios. Na avaliação geral do modelo, a análise de *deviance* (Tipo II, teste de Wald) retornou que a covariável TIPO DE VOGAL tem efeito significativo na duração ( $\chi^2 = 154,28$ ; gl = 1;  $p < 0,001$ ), bem como sua interação com a CONDIÇÃO experimental ( $\chi^2 = 4,62$ ; gl = 1;  $p = 0,031$ ). O efeito geral da CONDIÇÃO experimental também foi significativo ( $\chi^2 = 5,64$ ; gl = 1;  $p = 0,017$ ).

Esses resultados do modelo não eram esperados, pois, a partir da análise gráfica e das medidas de resumo, vemos que as médias tendem a ser muito próximas. A provável explicação para isso está relacionada ao efeito de uma das covariáveis do modelo, que testa  $H_0: \mu_1 = \mu_2 = \mu_k$ , e a rejeita se pelo menos um valor de  $\mu$  for diferente. Sendo assim, é necessário um teste de comparações múltiplas para identificar a origem da diferença. Neste caso, bastou avaliarmos o comportamento da duração das vogais nas duas condições experimentais. O resultado confirmou o esperado a partir da análise exploratória, isto é, as durações médias das vogais controle (pretônica) e a vogal resultante de sândi são estatisticamente iguais ( $\beta = -0,010$ ; EP = 0,036;  $z = -0,282$ ;  $p = 0,77$ ), confirmado pelo teste de Tukey. A origem da diferença está, portanto, nas durações de [i] e [u], mas não na sua posição dentro da frase fonológica, se Controle (vogal pretônica) ou Sândi (vogal parseada como vogal pretônica). Esse resultado faz sentido, uma vez que essas vogais possuem duração intrínseca estatisticamente diferente, sendo [i] mais curto que [u] (cf. Escudero et al., 2009; Figura 7 deste estudo).

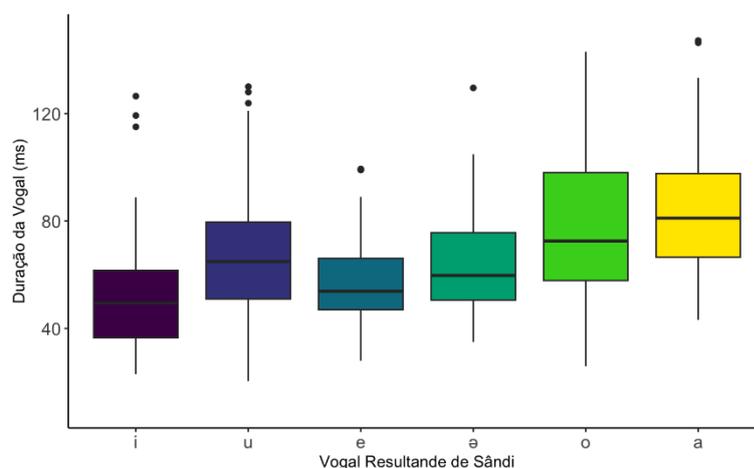


Fonte: Elaborado pelos autores.

Figura 7 - Duração das vogais [i] e [u] para cada condição experimental.

O fato de as vogais em Sândi e Controle apresentarem duração estatisticamente igual confirma a conclusão de Bisol (1992, 2000) de que a vogal resultante de SVE é sempre prosodicamente mais forte e, portanto, é um segmento associado a uma sílaba pretônica. A Figura 7 mostra as durações observadas na amostra, em que cada vogal não se diferencia entre as condições experimentais. Desse modo, permite-nos concluir que a elisão de uma vogal fonológica não tem consequências para redução de conteúdo fonético restante ou mesmo de um alongamento compensatório em  $V_2$ . Verifica-se que tanto [i] quanto [u] resultantes de sândi são semelhantes em duração aos contextos de sílaba pretônica. Tal fato parece favorecer a interpretação de que também não haja uma geminação intermediária em algum nível de processamento fonológico nos casos de  $V_1$  e  $V_2$  idênticas para uma degeminação no nível fonético, uma vez que a elisão é um fenômeno geral.<sup>1</sup>

Os resultados duracionais de monotongos em SVE são apresentados na Figura 8 e na Tabela 5, a seguir.



Fonte: Elaborado pelos autores.

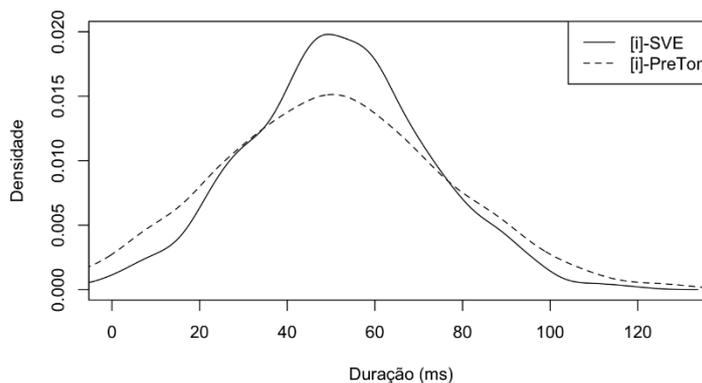
Figura 8: - Duração das vogais em SVE em escala logarítmica.

Tabela 5 - Média da duração (ms) das vogais em SVE e respectivos desvios-padrões.

	Média (ms)	Desvio-Padrão (ms)
i	52	22
e	57	20
ə	67	28
a	88	29
o	78	29
u	68	27

Fonte: Elaborado pelos autores.

Os resultados sobre a duração em SVE vão ao encontro do que Madruga (2017, p. 206) achou para as médias de vogais pretônicas do dialeto gaúcho: [i] = 49(26) ms, [e] = 69(32) ms, [a] = 70(19) ms, [o] = 72(31) ms e [u] = 53(26) ms. A diferença para este estudo refere-se à vogal central [ə], que encontramos em SVE e não na realização de pretônicas no PB, tal como se constata também no PE, cuja centralização de vogais ocorre em todas as posições átonas. Ao fazermos uma comparação entre os resultados encontrados neste trabalho com os reportados em Madruga (2017) simulados a partir dos parâmetros de uma distribuição probabilística normal para a duração das vogais, fica evidente que a duração das vogais tendem a ser estatisticamente idênticas.



Fonte: Elaborado pelos autores.

Figura 9: - Densidade da duração de [i] em SVE e [i] pretônico.

Para simular o comportamento da média nos dois estudos e compará-los, foram geradas duas amostras aleatórias normais de tamanho 1000 para as vogais [i] com os parâmetros de média e de variância amostrais obtidas neste trabalho e no de Madruga (2017), considerando que duração tem distribuição de probabilidade normal,  $X \sim N(\mu, \sigma^2)$ . A Figura 9 compara a densidade para a duração de [i] pretônico do dialeto gaúcho, ( $\bar{x} = 49$  ms,  $s = 26$  ms), e [i] resultante de SVE ( $\bar{x} = 52$  ms,  $s = 22$  ms). A partir da inspeção gráfica, vemos que o centro da densidade das duas curvas tende ao mesmo ponto, o que nos leva a considerar que ambas são vogais com o mesmo

comportamento. Logo, as vogais resultantes de sândi são vogais pretônicas, confirmando a proposta de Bisol (1992:84) de que as sílabas resultantes de SVE ficam sob o domínio do acento do vocábulo seguinte. Embora essa informação parece ser evidente do funcionamento de SVE, escapa, ao objeto desta análise, o status prosódico da estrutura resultante do sândi em termos de estruturação de outros níveis, tal como o parseamento das palavras prosódicas envolvidas. Caberá a trabalhos futuros discutir como se dá a formação das estruturas envolvidas no processo e da proeminências das sílabas acentuadas.

## 4 DISCUSSÃO

### 4.1 Os padrões de Sândi Vocálico Externo

Os resultados obtidos neste estudo sobre a elisão como SVE indicam que o fenômeno não está associado diretamente a fenômenos puramente fonéticos como taxa de elocução, isto é, rapidez da fala, tal como sugerem Meneses e Albano (2015). Ao contrário, os resultados evidenciam que a elisão, fenômeno geral, mantém-se constante nas repetições dos estímulos, enquanto apenas ditongação e coalescência parecem sofrer efeitos da rapidez da fala. A Figura 3 mostra relação inversa entre coalescência e ditongação: a primeira sofre um ligeiro aumento conforme se aumenta o número de repetições, enquanto a ditongação decresce drasticamente. Nesse sentido, pode-se hipotetizar que o falante, na maioria dos casos de SVE, opta pela elisão total de  $V_1$ , conforme se verificou na Tabela 2 (elisão: 52% dos casos). Por outro lado, nos outros 48% dos casos, de alguma forma, o falante parece evitar a elisão completa dos segmentos e mantém identidade parcial com o nível fonológico: ditonga ou coalesce totalmente os segmentos envolvidos (ditongação em 10% e coalescência em 38% dos casos). Como a COAL é um tipo de elisão de segmentos, pode-se supor que a preferência por elidir, total ou parcialmente (EL e COAL, respectivamente), corresponde a aproximadamente 90% dos casos.

Vê-se, na verdade, que ditongos [əj, əw] expressam a manutenção segmental, embora ocorram mudanças em suas características acústicas e uma reinterpretação no nível silábico, já que as vogais [i, u] em  $V_2$  tornam-se as semivogais [j, w]. A coalescência é um exemplo também de manutenção parcial dos segmentos fonológicos e não do tempo, já que a vogal resultante assimila as características de ambas vogais envolvidas na junção. Sendo assim, no contato de [a##u] a vogal de SVE é [o], enquanto na sequência [a##i] o resultado é a vogal média [e]. Em função desses resultados, a interpretação assumida é de que ocorre a seguinte sequência para a realização da coalescência.

(3)

#### **Sequência da Coalescência em SVE**

alçamento de /a/ > assimilação de posterioridade em ditongo > coalescência

A evidência de que os ditongos interagem com a coalescência é verificada no comportamento de /a/, que é alçada e assimilada em posterioridade pela vogal seguinte. Não se pode afirmar, no entanto, que a ditongação do tipo [ej] ou [ow] é uma fase necessariamente anterior à coalescência, mas que há apenas uma interação entre elas, já que o aumento de coalescências está associado com a redução de ditongos.

Portanto, o que verificamos é a existência de alternância entre ditongo e coalescência, tal como se propõe em (4).

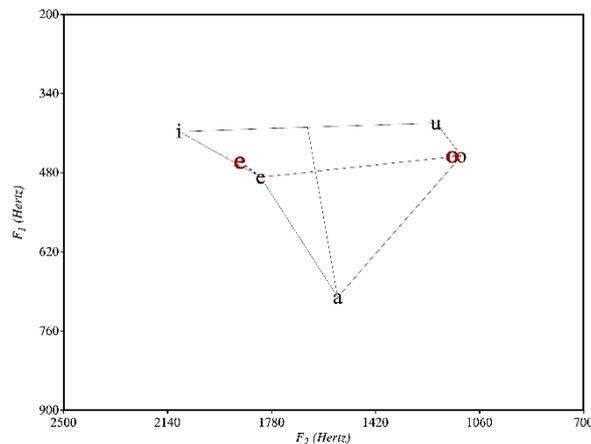
- (4)  
 Ditongação e Coalescência para /a ## i/  
 [əj] ~ [ej] ~ [e]  
 Ditongação e Coalescência para /a ## u /  
 [əw] ~ [ow] ~ [o]

A definição do mecanismo de coalescência observado nos nossos resultados e potencialmente presente na fala espontânea dos brasileiros é dada em (5).

(5)  
**Coalescência em SVE**

A vogal átona final /a/ de uma palavra fonológica e as vogais iniciais /i/ e /u/ átonas da palavra fonológica seguinte coalescem em altura e em posterioridade no domínio de uma frase fonológica.

Os resultados referentes às vogais coalescentes foram, então, comparados com valores disponíveis na literatura sobre vogais pretônicas. Na Figura 10, compara-se as vogais [e, o] resultantes de coalescência em SVE encontradas nesta pesquisa com o espaço acústico F1 x F2 a partir de valores disponíveis em Madruga et al. (2020).



Fonte: Elaborado pelos autores.

Figura 10 - Espaço acústico pretônico adaptado de Madruga et al. (2020) sobreposto às vogais [e, o] resultado de coalescência por SVE neste estudo.

Como se vê, os valores de F1 obtidos para [e] e [o] são muito próximos aos reportados por Madruga et al. (2020), Sandalo & Kenstowicz (2016), Madruga (2017), Barbosa et al. (2019). Em relação ao F1 das vogais pretônicas [e, o], Madruga et al. (2020) reportam F1 próximos a 411 Hz para [e] e 380 Hz para [o], enquanto Barbosa et al. (2019) reportam [e] com F1 médio de 466 Hz e [o] com F1 médio de 532Hz, valores semelhantes aos encontrados nesta pesquisa considerando os desvios-padrões.

#### 4.2 Proposta para Sândi Vocálico Externo no Português Brasileiro

O sândi externo no português, como processo de juntura de palavras, é entendido como um processo de ressilabação. Independentemente da variedade em que ele ocorra, o mecanismo fonológico processual é ressilabar os elementos de fronteira, muito embora seus resultados e bloqueios possam ser distintos em cada variedade da língua. Balduino (2022) relata a ocorrência dos mesmos processos conhecidos no PB no Português de São Tomé e Príncipe (PTSP), mas reconhece diferenças em termos das vogais envolvidas. Além disso, os mecanismos de bloqueio de SVE podem ser distintos no nível suprasegmental de ocorrência de sândi, tal como a fronteira do sintagma entoacional, que bloqueia o sândi no PE e não no PTSP e no PB (Tenani, 2002; Balduino, 2022).

Para sumarizar o panorama de SVE até então considerado na literatura da área, apresentamos no Quadro 1 as possibilidades combinatórias do encontro de vogais átonas e os processos de sândi resultantes do seu contato na frase fonológica, conforme as propostas de Bisol (1992, 1996, 2000 entre outros) para o fenômeno.

Quadro 1 - Resumo da proposta de Bisol (1992, 1993 entre outros) para SVE no PB.

V <sub>1</sub> ## V <sub>2</sub>		V <sub>2</sub>				
		i	e	a	o	u
V <sub>1</sub>	I	DEG	EL	DIT	DIT	DIT
	A	DIT	EL	DEG	EL	DIT
	U	DIT	DIT	DIT	EL	DEG

Fonte: Elaborado pelos autores.

O fato de V<sub>1</sub> ser átona final faz com que resulte foneticamente apenas as vogais [i, a, u]<sup>3</sup>, as quais podem ser seguidas pelas cinco vogais pretônicas na segunda palavra do sintagma. Percebe-se que, na juntura de vogais idênticas, a solução de Bisol é pela degeminação, e no encontro de [a] em V<sub>1</sub> com [i, u] em V<sub>2</sub> a língua pode operar com ditongação e variavelmente com a elisão (Bisol, 2000).

Passamos à discussão de cada caso detalhadamente. A elisão ocorre nas sequências [i ## e], [a ## e] e [a ## o] e [u ## o], enquanto para as sequências [a ## i] e [a ## u] a elisão alterna com ditongação decrescente e coalescência. Em relação aos ditongos, decorrem dessa proposta dois tipos: a) ditongação decrescente, para as sequências [a ## i] e [a ## u], únicas combinações possíveis para tal resolução silábica; b) ditongação crescente, casos em que V<sub>1</sub> é [i] ou [u], que possuem mais possibilidades combinatórias e podem sofrer ditongação segundo algumas restrições relativas às vogais nucleares.

Os ditongos decrescentes alternam com EL e são eles os alvos de COAL. Por outro lado, os ditongos crescentes resultantes de SVE obedecem às restrições de monotongação, conforme proposto por Madruga e Abaurre (2015). Os ditongos

<sup>3</sup> Não convém aqui fazer distinção entre [i, a, u] e [ɪ, ə, ʊ], pois essa diferença tem efeitos apenas na qualidade da vogal átona final de enunciado. Em sândi, interessa que resulta do alçamento das átonas finais apenas um sistema de vogais cardinais.

encontrados em sândi monotongam-se quando as vogais envolvidas concordam em posterioridade (Grupo I), como nos casos em (6).

(6)

Elisão de [i] e [u] em V<sub>1</sub> seguidos de /e/ e /o/ em V<sub>2</sub>

- a. i##e:  
 crime hediondo > crim[e]diondo \*crim[i]diondo  
 elite hegemônica > elit[e]gemônica \*elit[i]gemônica
- b. u##o:  
 falo horrores > fal[o]rrores \*fal[u]rrores  
 gosto objetivo > gost[o]bjetivo \*gost[u]bjetivo

Nesses casos, a vogal elidida é sempre V<sub>1</sub>, não podendo ser elidida V<sub>2</sub>, caso em que resultaria uma estrutura mal-formada. Através de teste simples de gramaticalidade em (6), pode-se dizer que SVE submete-se às restrições de monotongação, sendo a elisão o resultado ótimo. Em teste de gramaticalidade análogo, agora para efeito das consoantes precedentes a [a] ou [u] em V<sub>1</sub>, propomos que EL interage também com a palatalização de /t/ e /d/, bloqueando-a em dialetos que exibem palatalização dessas consoantes, como vemos em (7a-b), o que, por hipótese não ocorreria em dialetos não palatalizantes já não existiria processo a ser bloqueado. Esse tipo de interação de processos, provavelmente, funciona como pista da fronteira de palavras nos dialetos palatalizantes. Observa-se que só ocorre o bloqueio da palatalização se a vogal elidida for [a] ou [u], enquanto, se V<sub>1</sub> for [i], a palatalização ocorre (7c).

(7) Bloqueio de palatalização de /t, d/ em SVE (dialetos palatalizantes)

- a. meren[da ## i]scolar > meren[dis]colar  
 \*meren[dʒis]colar
- b. por[tu ## i]scuro > por[tis]curo \*por[tʃis]curo
- c. pare[di ## i]sverdeada > pare[dʒis]verdada

Bisol (1992) exemplifica que em <gente esquece> gen[tʃi]squece, ocorre a palatalização, porém, para a autora, esse é um caso de DEG<sup>4</sup>. Já ditongação realiza-se nos casos em que a fonotaxe de V<sub>1</sub>##V<sub>2</sub> não é suscetível à elisão, formando-se apenas ditongos crescentes discordantes em posterioridade (ditongos de Grupo II), conforme proposta de Madruga e Abaurre (2015).

(8) Ditongação de [i] e [u] em V<sub>1</sub> seguidos de /e/ e /o/ em V<sub>2</sub>

- c. i##o:  
 fase horrorosa > fa[zjo]rrososa  
 \*fas[o]rrososa  
 leite homogêneo > lei[tʃjo]mogêneo \*leit[o]mogêneo
- d. u##e:  
 direito eficaz > direi[twe]ficaz ?direit[e]ficaz  
 posto eficiente > pos[twe]ficiente ?post[e]ficiente

<sup>4</sup> Não discutimos o estatuto fonológico da vogal final, se /e/ ou /i/ e as consequências que assumir uma ou outra vogal na subjacência traz para a implementação da palatalização. Para os efeitos desejados neste estudo, basta atentarmos para o funcionamento do bloqueio quando a vogal final a ser elidida é [a] ou [u].

Considerando essa discussão, propõe-se que o português brasileiro apresenta apenas dois principais processos de sândi vocálico externo: elisão e a ditongação crescente. O terceiro, a coalescência, é entendido como um subprocesso da elisão, já que é um tipo de elisão parcial na qual se apaga o tempo fonológico de V<sub>1</sub> e não o seu conteúdo melódico inteiramente. A evidência é que a coalescência sempre alterna com a elisão nas duas possibilidades combinatórias, como evidenciamos no Quadro 2. Ainda, consideramos que o fenômeno de degeminação proposto por Bisol (1992, 1996) constitui essencialmente uma elisão da vogal átona final e, portanto, assumimos que não há degeminação, mas apenas elisão. O Quadro 2 resume a nova proposta para o sistema de SVE para o PB.

Quadro 2 - Proposta de padrões de SVE no português brasileiro.

V <sub>1</sub> ## V <sub>2</sub>		V <sub>2</sub>				
		i	e	a	o	u
V <sub>1</sub>	i	EL	EL	DIT'	DIT'	DIT'
	a	EL ~ DIT' ~ COAL	EL	EL	EL	EL ~ DIT' ~ COAL
	u	DIT'	DIT'	DIT'	EL	EL

Fonte: Elaborado pelos autores.

As evidências linguísticas da preferência pela elisão em contextos de segmentos adjacentes idênticos pode ser encontrada no funcionamento do sândi consonantal em PB. Vejamos o que acontece com /s/ final. Temos que a consoante final /s/ é rressilabificada quando seguida por vogal ou assimilada em vozeamento quando seguida por consoante vozeada, tal como em (9).

(9) Assimilação de Vozeamento

- a. [s ## C<sub>[α voz]</sub>]:
- |                |   |                  |
|----------------|---|------------------|
| copos pequenos | > | copo[s.p]equenos |
| braços bonitos | > | braço[z.b]onitos |

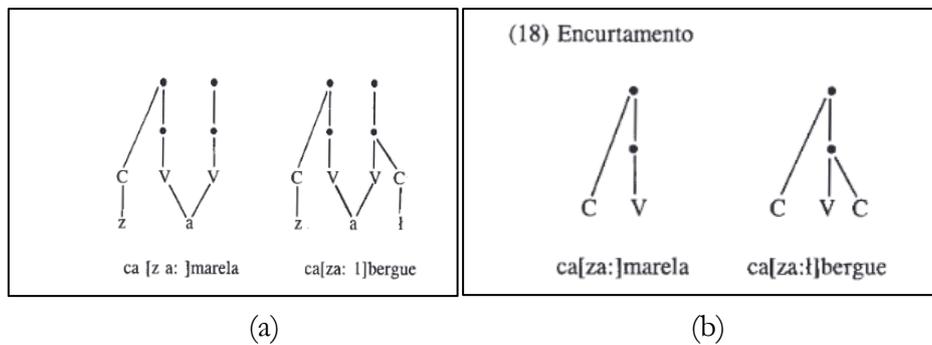
O processo de assimilação de vozeamento não ocorre, entretanto, quando C<sub>2</sub> é fricativa coronal, seja anterior ou não. Se seguissemos a mesma lógica do que se propõe para vogais, poderíamos propor a existência de uma degeminação consonantal nos casos de [s ## s] (veja exemplos em 10a). Todavia, ao empregarmos um teste de comutação de segmentos e fixarmos uma consoante coronal [-anterior] em C<sub>2</sub>, o resultado de sândi é a elisão da consoante /s/ em C<sub>1</sub>, como em (9b-c).

(10)

- a. [s ## s]
- |              |   |              |
|--------------|---|--------------|
| roupas secas | > | roupa[s]ecas |
| boas sopas   | > | boa[s]opas   |
- b. [s ## z]
- |                |   |                |                 |
|----------------|---|----------------|-----------------|
| lindas zebras  | > | linda[z]ebras  | *linda[s]ebras  |
| contas zeradas | > | conta[z]eradas | *conta[s]eradas |

- c. [s ## ]  
 casas chiques > casa[ʃ]iques \*casa[s]iques  
 duas chaves > dua[ʃ]aves \*dua[s]aves
- d. [s ## ʒ]  
 dois jovens > doi[ʒ]ovens \*doi[z]ovens  
 bons jogos > bon[ʒ]ogos \*bon[z]ogos

Em (10b-d), vemos que a manutenção de C<sub>1</sub> resultaria em uma estrutura mal-formada, evidenciando, portanto, que os casos de sequências segmentais idênticas em contexto de sândi externo sofrem elisão em nível abstrato intermediário, uma vez que, no nível fonológico, há uma fricativa do tipo /S/ em coda da primeira palavra do sintagma, que não chega ao nível fonético, sendo apagada no momento da produção. O resultado em (10a) é consequentemente uma opacidade fonológica, uma vez que a produção fonética não fornece evidência suficiente para se determinar qual é o segmento resultante, se C<sub>1</sub> ou C<sub>2</sub> e o mesmo para V<sub>1</sub> ou V<sub>2</sub>. Essa evidência só pode ser capturada a partir da análise do comportamento do sândi consonantal para sequências de coronais (10b-d), em que poderia ocorrer uma assimilação de vozeamento, se /s/ for seguido por [ʒ], ou ser realizado como [s], se [ʃ] for C<sub>2</sub>. Como não é o que ocorre, esse comportamento revela que [s, z, ʃ, ʒ] formam classe natural, sendo o grupo tratado como constituído de segmentos idênticos para sândi externo. Sendo assim, faz sentido que o comportamento das consoantes nos revele o que ocorre em SVE, já que a única diferença dos sândis é o tipo de segmento envolvido. Para Bisol (1992), esse processo envolve o Princípio do Contorno Obrigatório (OCP), o qual opera ao encontrar segmentos adjacentes idênticos, fundindo as duas vogais no nível melódico e, num primeiro momento da derivação, resultando um elemento associado a duas unidades temporais. No final do processo, ocorre o alongamento da vogal, a qual será encurtada em seguida para se gerar o resultado esperado por OCP, como na Figura 11.



Fonte: Bisol (1992, p. 92).

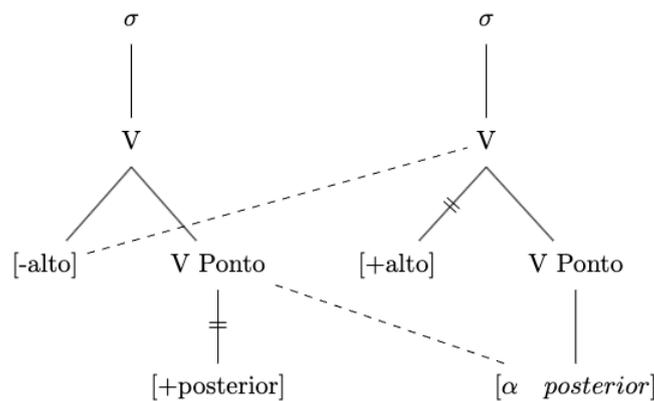
Figura 11 - Fases da derivação da degeminação para V<sub>1</sub> e V<sub>2</sub> idênticas em sândi.

Ao eliminarmos o processo de degeminação como uma possibilidade de SVE, damos uma opção analítica mais econômica, pois o interpretamos como elisão e, por consequência, o processo fica condicionado aos mecanismos operacionais da elisão. Os mecanismos que governam SVE ficam mais simples e determinados por apenas duas principais operações de ressilabificação nos níveis abstratos, a saber: a que elide vogais e a operação que torna glide as vogais altas, formando ditongos. Na proposta deste trabalho, postulamos que não há degeminação e que este mecanismo de encurtamento não é necessário porque a V<sub>1</sub> é dissociada assim que encontra uma V<sub>2</sub> idêntica na fronteira vocálica. Sendo assim, a vogal é dissociada do núcleo silábico e

prontamente elidida, enquanto a consoante do *onset* é associada à sílaba seguinte por ressilabificação, tal como propõe Bisol (1992: 98). DEG, então, não é necessária na derivação, ocorrendo em seu lugar EL.

Formalmente, a coalescência é resultado da fusão de [-alto] e [α posterior], situação em que emerge [e, o]. Entende-se a coalescência como um subprocesso da elisão porque o tempo fonológico de V<sub>1</sub> é apagado, como já demonstramos na seção 3.3. No entanto, em fase anterior ao apagamento as vogais se fundem no nível melódico e resulta apenas o tempo do segmento a ser apagado na derivação. A coalescência resultante é assimilatória do tipo redutiva (cf. Zaleska, 2020, p. 699), cuja formalização para o PB é dada em (11). A formalização fornecida aqui está fundamentada em um tipo de representação autosegmental, versão ‘Coloured Containment’ (van Oostendorp 2006, 2007), uma vez que a proposta sobre o funcionamento de SVE aqui empreendida não se pretende como um revisionismo total das principais propostas representacionais (cf. Bisol, 1993, 2000). No processo de coalescência redutiva, tem-se uma redução, na superfície, do número de segmentos subjacentes, ao não escandir nós, mas se espalha o conteúdo abaixo desses aos segmentos escandidos.

(11)



Ademais, duas evidências adicionais podem ser trazidas à argumentação no que concerne à coalescência e à exclusão de DEG: i) a emergência de [e] em SVE é largamente aceita na pronúncia de numerais, como em (12); ii) os contextos de bloqueio de EL são os mesmos de DEG.

(12) Coalescência em Φ numerais

- |    |                      |                 |   |
|----|----------------------|-----------------|---|
| a. | setenta e cinco (75) | seten[tej]cinco | ~ |
|    |                      | seten[te]cinco  |   |
| b. | trinta e dois (32)   | trin[tej]dois   | ~ |
|    |                      | trin[te]dois    |   |

A pronúncia coalescente de numerais como os exemplificados parece ser bem aceita pelos falantes e, se assim for, seria uma evidência do funcionamento da coalescência em SVE em variedades do PB. A variação entre ditongos assimilados e vogais coalescentes em (12) demonstra que a coalescência no PB é assimilatória

reduzitiva, tal como propõe Zaleska (2020) e mostra que o fenômeno é produtivo na língua e não apenas um resultado experimental.

Ainda, nos bloqueios de DEG e EL, Bisol (1993, 19:22) afirma que nos casos de  $V_1 \#\#V_2$  ocorre os bloqueios de EL, mas não de DEG (exemplo 2b). No entanto, Tenani (2005), ao investigar a interação de domínios prosódicos com o acento nuclear, demonstra que o bloqueio de DEG e o de EL estão associados à coincidência do acento em uma das vogais e ao acento nuclear. Para Tenani (2005:19), “*não há dúvidas da relevância de  $\Phi$  quer para o bloqueio da DG e da EL, quer para o bloqueio da resolução de choque de acentos*”. Desse modo, vemos que as restrições sob DEG e EL são as mesmas porque o fenômeno é o mesmo, EL.

## 5 CONSIDERAÇÕES FINAIS

Este trabalho procurou compreender o funcionamento do SVE a partir de dados experimentais. Os resultados reportados referem-se ao tipo de processo de SVE, à qualidade das vogais resultantes e ao parseamento prosódico da sílaba resultante de sândi. Nossos achados permitiram trazer à discussão uma nova proposta de análise fonológica do fenômeno, que descrevemos em § 4.

Em suma, temos que da fronteira  $V_1 \#\# V_2$  surgem três padrões: a elisão, a coalescência e a ditongação. O fenômeno preferencial é a elisão de  $V_1$ , cuja ocorrência corresponde a mais de 50% dos casos, seguido pela coalescência com 37,4%. Os ditongos que interagem com EL são sempre decrescentes em função de restrição fonotática dada pela língua para algumas sequências de vogais e realizaram-se com alçamento de [a] para [ə] ou foram completamente assimilados para [e] ou [o], dependendo da posterioridade de  $V_2$ . Esse resultado evidenciou que o ditongo decrescente assimilado está relacionado com a coalescência, sendo esta última uma assimilação redutiva do primeiro. COAL é, portanto, um subprocesso governado por EL, cujos mecanismos que atuam para garantir formas elididas parecem atuar de maneira mais preponderante na estrutura da fronteira de palavras.

A coalescência é um achado inédito na descrição do processo no português brasileiro e, a partir desses achados experimentais e dos exemplos disponíveis ao explorarmos a combinatória fonológica, propomos uma nova interpretação dos resultados de sândi vocálico externo para o PB, que foi detalhadamente apresentada nas seções 3 e 4. O resultado do sistema em SVE é mais econômico formalmente e em termos de processo porque se ajusta ao encontrado no uso real da língua e porque elimina processos derivacionais de alongamento de vogais que só existiriam no plano abstrato, os quais nunca chegam à superfície. O funcionamento do sândi consonantal entre fricativas coronais, discutido em (10), auxiliou-nos a elucidar o que acontece com vogais idênticas, evidenciando a preferência por apagamento em posições finais e preservando as iniciais. Como consequência, esses fatos nos levam a postular que DEG é EL em condições segmentais idênticas.

Além disso, vimos também que o processo de ressilabação do SVE confirma a proposta de Bisol (1992) sobre o seu estatuto prosódico no domínio do acento da palavra à direita. Isto é, a sílaba resultante de SVE realiza-se como sílaba pretônica. A diferença entre sílabas em SVE e sílabas pretônicas plenas refere-se ao sistema de seis vogais fonéticas em SVE, enquanto a sílaba pretônica de uma palavra fonológica apresenta cinco vogais se não sofrer nenhum processo assimilatório. Outra diferença

trazida à discussão é a ocorrência das consoantes [t, d] não-palatalizadas, o que não ocorre no interior de palavra fonológica se houver um /i/ fonológico.

De modo geral, várias questões ainda persistem para o futuro desta pesquisa, como investigar os casos de bloqueio de SVE, suas consequências acústicas e a categorização fonológica feita pelos falantes. Faz-se necessária também um detalhamento da relação dos ditongos decrescentes e a coalescência observada neste trabalho. Além de investigar esses pontos, vale, num desdobramento, analisar dados de fala espontânea e até semi-espontânea para mapearmos a ocorrência do fenômeno no uso da língua, bem como na sua abrangência dialetal. De todo modo, este estudo pretende contribuir para a análise de SVE, de modo que traz não apenas resultados inéditos, mas também uma proposta de interpretação fonológica que, salvos os erros analíticos, inova em relação ao já proposto na literatura da área.

### AGRADECIMENTOS

Agradecemos às Profas. Bernadete Abaurre e Amanda Balduino, bem como ao aluno Gabriel Castelano Millas pela leitura cuidadosa e pela discussão de versões anteriores deste texto. Agradecemos também à FAPEMIG (Edital 04/2022, UFMG) pela Bolsa de Iniciação Científica, atribuída ao segundo autor, que permitiu a finalização da pesquisa.

### REFERÊNCIAS

- Abaurre MBM. Acento frasal e processos fonológicos segmentais. *Letras de Hoje*. 1996;31(2). Disponível em: <https://revistaseletronicas.pucrs.br/fale/article/view/15591>.
- Albano EC. O português brasileiro e as controvérsias da fonética atual: pelo aperfeiçoamento da fonologia articulatória. *DELTA*. 1999;15:23-50. Disponível em: <https://doi.org/10.1590/S0102-44501999000300002>.
- Albano EC. O gesto e suas bordas: esboço de fonologia acústico-articulatória do português brasileiro. Campinas: ALB; 2001.
- Balduino AM. Fonologia do português de São Tomé e Príncipe [tese]. São Paulo: Universidade de São Paulo; 2022. Disponível em: <https://repositorio.usp.br/item/003104817>.
- Barbosa PA, Finardi A, Silva AP, Lucente L. Harmonia vocálica e coarticulação vogal a vogal em duas variedades do português brasileiro. *DELTA*. 2019;35. Disponível em: <https://doi.org/10.1590/1678-460X2019350202>.
- Bisol L. A elisão, uma regra variável. *Letras de Hoje*. 2000;35(1):319-30. Disponível em: <https://revistaseletronicas.pucrs.br/fale/article/view/14770>
- Bisol L. O sândi e a ressilabação. *Letras de Hoje*. 1996;31(2):159-68. Disponível em: <https://revistaseletronicas.pucrs.br/fale/article/view/15601>
- Bisol L. Sândi externo: o processo e a variação. In: Kato M, organizadora. Gramática do português falado: convergências. Campinas: Unicamp; 1994. Vol. V. p. 55-96.
- Bisol L. Sândi vocálico externo. In: Ilari R, organizador. Gramática do português falado II. Campinas: Editora da Unicamp; 1993. p. 21-38.
- Bisol L. Sândi vocálico externo: degeminação e elisão. *Caderno de Estudos Linguísticos*. 1992;23:83-101. Disponível em: <https://doi.org/10.20396/cel.v23i0.8636847>
- Boersma P, Weenink D. Praat: doing phonetics by computer [programa de computador]. Versão 6.6.22. 2022. Disponível em: <http://www.praat.org/>.

- Fant G. Acoustic theory of speech production. The Hague: Mouton; 1960. DOI: <https://doi.org/10.1515/9783110873429>.
- Flemming E. The phonetics of schwa vowels. Phonological weakness in English. 2009;493:78-95. Disponível em: <https://web.mit.edu/flemming/www/paper/schwaphonetics.pdf>.
- Herslund M. Portuguese sandhi phenomena. In: Andersen H, editor. Sandhi phenomena in the languages of Europe. Berlin; New York; Amsterdam: Mouton de Gruyter; 1986. Disponível em: <https://doi.org/10.1515/9783110858532>
- Kenstowicz M., Sandalo F. Pretonic vowel reduction in Brazilian Portuguese: harmony and dispersion. J Port Linguist. 2016;15(6). Disponível em: <https://jpl.letras.ulisboa.pt/article/id/5642/>.
- Madruga MR. The phonetics and phonology of Brazilian Portuguese [ATR] harmony: a fonética e fonologia da harmonia de [ATR] no português brasileiro [tese]. Campinas: Universidade Estadual de Campinas; 2017. Disponível em: <https://doi.org/10.47749/T/UNICAMP.2017.995493>.
- Madruga MR, Abaurre MBM. Restrições fonotáticas de onset e ditongos crescentes em português. Rev. ABRALIN. 2015;14(1):339-77. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/1244>.
- Madruga MR, Hamann S, Abaurre MBM. Gradient and categorical assimilation of pretonic vowels in Brazilian Portuguese. J Port Linguist. 2020;19(1):12. DOI: <https://doi.org/10.5334/jpl.234>.
- Meneses F. Uma visão dinâmica dos processos de apagamento de vogais no português brasileiro [tese]. Campinas: Universidade Estadual de Campinas; 2017. Disponível em: <https://repositorio.unicamp.br/acervo/detalhe/976203>.
- Nogueira MV. Aspectos segmentais dos processos de sândi vocálico externo no falar de São Paulo [dissertação]. São Paulo: Universidade de São Paulo; 2007. Disponível em: <https://teses.usp.br/teses/disponiveis/8/8139/tde-12112007-150159/es.php>.
- Paulino NMR. Fenómenos de sândi vocálico em variedades do português europeu [dissertação]. Lisboa: Universidade de Lisboa; 2016. Disponível em: <http://hdl.handle.net/10451/26025>.
- R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2023. Disponível em: <https://www.R-project.org/>.
- Tenani L. Considerações sobre a relação entre processos de sândi e ritmo. Estud Lingua(gem). 2006;5(1):105-22. Disponível em: <https://periodicos2.uesb.br/index.php/estudosdalinguagem/article/view/1011>.
- Tenani L. Domínios prosódicos no português do Brasil: implicações para a prosódia e para a aplicação de processos fonológicos [tese]. Campinas: Universidade Estadual de Campinas; 2002. Disponível em: <https://doi.org/10.47749/T/UNICAMP.2002.253138>.
- Tenani L. O bloqueio do sândi vocálico em PB e em PE: evidências da frase fonológica. Organon. 2004;18(36). Disponível em: <https://seer.ufrgs.br/index.php/organon/article/view/31151>.
- Van Oostendorp M. A theory of morphosyntactic colours [manuscript]. Meertens Institute; 2006. Disponível em: [https://www.researchgate.net/profile/Marc\\_Oostendorp](https://www.researchgate.net/profile/Marc_Oostendorp).
- Van Oostendorp M. Derived environment effects and consistency of exponence. Freedom of analysis. 2007;95:123-48. DOI: [10.1515/9783110198591.123](https://doi.org/10.1515/9783110198591.123).
- Vianna P. Sândi vocálico externo: o processo e a variação na cidade de Florianópolis - SC [dissertação]. Porto Alegre: Universidade Federal do Rio Grande do Sul; 2009. Disponível em: <https://lume.ufrgs.br/handle/10183/18353>.
- Zaleska J. Coalescence as autosegmental spreading and delinking. Phonology. 2020;37(4):697-735. DOI: <https://doi.org/10.1017/S0952675720000317>.

## Vogais tônicas orais do kabuverdianu do Príncipe: uma análise fonética

### *Oral stressed vowels of the Príncipe Kabuverdianu: a phonetic analysis*

Shirley Freitas\*

*Universidade da Integração Internacional da Lusofonia Afro-Brasileira;*  
São Francisco do Conde, BA, Brasil

Amanda Macedo Balduino\*\*

*Universidade Estadual de Campinas; Campinas, SP, Brasil*

**Resumo:** Com a migração caboverdiana para São Tomé e Príncipe, iniciada em 1903, o kabuverdianu foi levado para o Golfo da Guiné e passou a fazer parte da ecologia linguística da região. Considerando esse cenário, este estudo se debruça sobre o kabuverdianu da ilha do Príncipe (KVP), apresentando uma descrição fonética acústica do primeiro (F1) e segundo formantes (F2) e da duração que caracterizam as vogais orais tônicas nessa língua. Como corpora, utilizamos dados coletados no Príncipe em 2018 com quatro falantes de kabuverdianu como língua materna (Bandeira; Freitas, 2018). Os dados foram analisados com auxílio do programa Praat (Boersma; Weenick, 2022). Em posição tônica, foram encontrados nove fones vocálicos orais ([i, e, ε, ə, ɐ, a, ɔ, o, u]), ampliando as descrições do kabuverdianu de Santiago que mencionavam somente 2 vogais centrais. O comportamento das vogais médias e centrais sugere que a dispersão acústica é relevante na definição das vogais da língua. As vogais médias parecem estar mais próximas entre si (com a diferença entre F1 sendo de 75 Hz para as coronais e 98 Hz para as dorsais). Esse fato aproxima acusticamente as vogais médias-altas e médias-baixas, uma vez que os dois pares ([e] e [ε], [o] e [ɔ]) passam a ficar menos dispersos no espaço acústico.

**Palavras-chave:** Kabuverdianu. São Tomé e Príncipe. Vogais tônicas orais. Descrição acústica.

**Abstract:** With the Cape Verdean migration to São Tomé and Príncipe, which began in 1903, Kabuverdianu was taken to the Gulf of Guinea and became part of the region's linguistic ecology. Considering this scenario, this study focuses on Kabuverdianu from the island of Príncipe (KVP), presenting an acoustic phonetic description of the first (F1) and second formants (F2) and the duration that characterize the oral stressed vowels in this language. As corpora, we use data collected in Príncipe in 2018 with four speakers of Kabuverdianu as their mother tongue (Bandeira; Freitas, 2018). The data were analyzed using the Praat program (Boersma; Weenick, 2022). In tonic position, nine oral vowel phones were found ([i, e, ε, ə, ɐ, a, ɔ, o, u]), expanding the descriptions of Santiago's kabuverdianu that mentioned only 2 central vowels. The behavior of mid vowels and central vowels suggests that acoustic dispersion is relevant in defining the language's vowels. The middle vowels seem to be closer to each other (with the difference between F1 being 75 Hz for the front and 98 Hz for the back). This fact acoustically brings

---

\* Docente, Instituto de Humanidades e Letras, Universidade da Integração Internacional da Lusofonia Afro-Brasileira, São Francisco do Conde, BA, Brasil; [shirleyfreitas@unilab.edu.br](mailto:shirleyfreitas@unilab.edu.br)

\*\* Docente, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, SP, Brasil; [amandamb@unicamp.br](mailto:amandamb@unicamp.br)

together the mid-high and mid-low vowels, since the two pairs ([e] and [ɛ], [o] and [ɔ]) become less dispersed in the acoustic space.

**Keywords:** Kabuverdianu. São Tomé and Príncipe. Oral stressed vowels. Acoustic description.

## 1 INTRODUÇÃO

São Tomé e Príncipe (doravante STP) é um país da África Ocidental que possui uma realidade multilíngue, com a presença de línguas crioulas autóctones de base portuguesa – santome (ISO 639-3: cri), lung'le (ISO 639-3: pre) e angolár (ISO 639-3: aoa) –, do português (em suas diversas variedades) e do kabuverdianu (ISO 639-3: kea). Dentre essas diversas línguas, este estudo enfoca o kabuverdianu, língua oriunda do contato entre o português dos séculos XVI e XVII e línguas africanas (sobretudo dos grupos mandê e atlântico) que começou a se desenvolver em Cabo Verde no início do século XVI e hoje é o idioma mais difundido no arquipélago (Quint, 2000). O kabuverdiano foi transplantado para o Golfo da Guiné no início do século XX e, atualmente, é falado em STP. Mais precisamente este artigo trata do kabuverdianu do Príncipe (doravante KVP), ilha que concentra a maior parte dos caboverdianos e seus descendentes (Carreira, 1983; Feio, 2016), apresentando as vogais fonéticas tônicas dessa variedade a partir de uma descrição acústica, que considera fatores como o primeiro e o segundo formantes em Hertz e a duração em milissegundos.

A justificativa para este trabalho se encontra no fato de que, apesar de a comunidade caboverdiana ser bastante expressiva em STP, sobretudo no Príncipe, a variedade de kabuverdianu falada no país ainda não tem sido tema de estudos sistemáticos – com exceção dos trabalhos de Freitas et al. (2021) e Freitas (em preparação). Na verdade, grande parte dos estudos do kabuverdianu se concentra nas variedades faladas em Cabo Verde, especialmente em Santiago, capital do país, de onde se iniciou o povoamento das demais ilhas e que possui a variedade dialetal mais antiga e conservadora (Quint, 2000). Assim, com este estudo, será possível contribuir com a compreensão da complexa ecologia linguística do Príncipe, ampliando, ainda, a literatura dedicada ao kabuverdianu.

O artigo está organizado da seguinte forma: a seção 2 discute brevemente o contexto sociohistórico da migração caboverdiana para STP, enfocando seus resultados linguísticos. A seção 3 se debruça sobre a metodologia da pesquisa: o corpus analisado, os participantes da pesquisa e os procedimentos de análise. Na seção 4, encontra-se a discussão das vogais tônicas orais do kabuverdianu do Príncipe, apresentando inicialmente os resultados referentes ao primeiro e ao segundo formantes, para em seguida discutir brevemente os valores da duração. Por fim, a seção 5 apresenta as considerações finais da pesquisa, apontando seus desdobramentos futuros.

## 2 MIGRAÇÃO CABOVERDIANA PARA STP

Remontando ao final do século XV, a colonização portuguesa na ilha que viria a ser STP foi iniciada a partir da implantação de um entreposto de escravizados e posteriormente mediante o estabelecimento de um centro monocultor açucareiro (cf. Hagemeyer, 2009; Bandeira, 2017), atividade econômica que durou até o final do século XVI. Neste período, a economia entra em declínio devido a dificuldades

climáticas referentes à fabricação do açúcar, o que se refletia em um produto de menor qualidade quando comparado ao de outras regiões como o Brasil, ataques de navios franceses, sucessivas pragas que atingiram as plantações, entre outros fatores.

Em meados de 1800 e depois de uma crise que durou alguns anos, o arquipélago é recolonizado, um período caracterizado pela monocultura de café e depois de cacau. Como consequência a esse fato, temos o estabelecimento das roças, propriedades monocultoras de pequeno, médio e grande porte que, atualmente, constituem hotéis de luxo. Segundo Seibert (2015), algumas características dessa nova fase são um maior contingente de colonos portugueses e a introdução de um novo segmento social, os trabalhadores contratados ou serviçais. O trabalho contratado está intrincado à abolição da escravização em 1875<sup>1</sup>, a qual foi marcada, diante das condições precárias de trabalho, pela recusa de forros, angolares e principenses, população nativa de STP, de trabalhar nas roças. Em decorrência disso, um grande contingente de mão de obra passou a ser requisitado, entre o começo de 1900 e 1922, de outras ex-colônias portuguesas como Angola e Moçambique e, a partir de 1920, de Cabo Verde (Carreira, 1983).

No início do século XX (1902/1903), os caboverdianos começam a chegar em STP e nos anos seguintes a imigração aumenta. Segundo Eyzaguirre (1986), entre os anos de 1903 e 1928, vieram a STP 6.305 contratados caboverdianos, sendo 1903 e 1921 os anos com maiores chegadas: 1.083 e 2.707, respectivamente. Os dados do autor mostram ainda que, nesse período, as chegadas mais expressivas eram de trabalhadores angolanos (50.444 imigrantes) e moçambicanos (43.053 imigrantes). De acordo com Eyzaguirre (1986), as chegadas de caboverdianos em STP vão se tornando cada vez mais significativas, tanto que, após 1945, a maior parte dos trabalhadores contratados passa a ser oriunda de Cabo Verde.

De acordo com Carreira (1983), entre 1902 e 1970, 81.059 caboverdianos foram para STP, estando incluídas nas cifras tanto casos de migração voluntária quanto forçada. Ainda com relação às cifras, Carreira (1983) menciona que a maior parte dos caboverdianos migrou para o Príncipe, enquanto em São Tomé predominaram trabalhadores angolanos e moçambicanos. O autor apresenta dados do Anuário de Lisboa de 1917, que aponta os seguintes números da migração caboverdiana para STP: (i) em 1915, 58 caboverdianos foram para São Tomé e 907 para o Príncipe; (ii) em 1916, 86 caboverdianos chegaram em São Tomé e 677 no Príncipe; (iii) em 1917, o número de migrantes caboverdianos que migrou para São Tomé foi de 43 e para o Príncipe 190. Contudo, nem todos os registros consultados fazem distinção entre as duas ilhas, colocando em muitos casos (como no período de 1941 a 1949), o destino de todos os contratados como sendo São Tomé. Uma das possíveis justificativas, segundo o autor, poderia estar no fato de os navios primeiro chegarem em São Tomé e só em seguida partirem para o Príncipe.

Além de uma motivação mais aventureira, de desbravar novas terras, autores como Nascimento (2008) apontam a fome, a miséria e a pobreza decorrentes dos sucessivos episódios de seca em Cabo Verde como as principais motivações para a migração. Diferentemente dos angolanos e moçambicanos, muitos contratados caboverdianos migraram junto com suas famílias, o que contribuiu largamente para a permanência desse segmento em STP após o fim dos contratos, sobretudo no Príncipe.

---

<sup>1</sup> 1878 de acordo com Nascimento (2010).

Quanto à procedência principal dos contratados, grande parte deles partiu de Santiago (Carreira, 1983), razão pela qual a comparação neste artigo focará a variedade falada nesta ilha. Na verdade, apesar de se usar o rótulo de trabalho contratado, com um contrato estipulado de 5 anos em média e o compromisso de pagamento de salários, as condições de trabalho replicavam a escravidão, com a prática de castigos físicos, do toque de recolher e a impossibilidade de livre circulação nas ilhas (Eyzaguirre, 1986; Nascimento, 2004, 2007, 2008; Seibert, 2015; Semedo, 2016).

Com relação aos aspectos linguísticos da comunidade caboverdiana, junto ao português – falado por 98,4% da população de STP de acordo com o censo (INE, 2012) – encontra-se também o kabuverdianu, sendo o domínio dessa língua considerado como um indicativo do pertencimento à comunidade caboverdiana, segundo Semedo (2016). Desse modo, conjectura-se que a maior parte dos falantes de kabuverdianu é bilíngue<sup>2</sup> em kabuverdianu e português.

Agostinho (2015) afirma que o kabuverdianu possui vitalidade no Príncipe, havendo mais falantes desta língua do que de lung'Ie, língua nativa da ilha. A autora afirma ainda haver falantes monolíngues de kabuverdianu, sobretudo em asilos no interior do Príncipe. Os dados do Censo de 2012 confirmam esse maior uso do KVP: em STP, 14.725 pessoas declararam ser falantes de kabuverdianu, o que equivale a 8,5% da população, ao passo que 1.760 afirmaram falar lung'Ie, correspondendo a 1% do total<sup>3</sup>.

Além de estar em contato com as línguas crioulas locais (santome e angolar em São Tomé e lung'Ie no Príncipe), em STP, o kabuverdianu também se relaciona com diferentes variedades de português. Balduino (2022) aponta que a difusão do português no cenário multilíngue do arquipélago com o posterior incremento no seu uso, chegando a ser a língua mais falada, culminou no surgimento de variedades próprias de português, como o português de São Tomé (PST) e o português do Príncipe (PP), que fazem parte de uma macrovariedade urbana, o português de São Tomé e Príncipe (PSTP). Além dessas variedades locais, segundo Santiago e Balduino (2023), o português europeu (PE) continua tendo uma presença política marcante no arquipélago, sendo a referência usada na documentação oficial e alvo do ensino, ainda que professores e estudantes não sejam usuários dessa norma, desconsiderando-se as variedades locais. Em vista desse cenário, na subseção 4, será realizada a comparação entre o KVP e o PST, o PP e o PE, dada a convivência entre as línguas, buscando-se observar ainda se o KVP possui mais similaridades com o PP do que com o PST, dado que, no caso do PP, há um contato direto.

Após essa breve caracterização sociohistórica da migração caboverdiana para STP, a seção 3 trata da metodologia da pesquisa.

---

<sup>2</sup> Neste artigo, o termo bilinguismo é usado para se referir ao domínio de duas línguas, sendo ambas usadas para comunicação.

<sup>3</sup> O Censo não traz dados separados por ilha. Vale mencionar, por fim, que o Censo também não discrimina se a língua é falada como língua materna ou não, tampouco refere as situações linguísticas em que os idiomas são empregados.

### 3 METODOLOGIA

Esta seção é dedicada à metodologia. Para tanto, explicitamos a constituição do corpus, a caracterização dos participantes e, por fim, os procedimentos de análise.

#### 3.1 Corpus e Trabalho de Campo

O corpus desta pesquisa é constituído por 523 ocorrências de 256 itens lexicais coletados em trabalho de campo com quatro falantes, na ilha do Príncipe, em 2018 (Bandeira; Freitas, 2018). A coleta das palavras foi baseada na lista Swadesh (Graham; Graham, 2014) e privilegiou itens do vocabulário básico<sup>4</sup> do KVP, visto que o propósito central do trabalho de campo era a documentação e descrição linguística dessa língua. Assim sendo, tendo em vista que o KVP é uma variedade ainda não descrita do kabuverdianu, o trabalho de campo realizado em 2018 buscou coletar dados *in loco*, dentro da comunidade de fala de caboverdianos na ilha do Príncipe, de modo que tal variedade pudesse ser investigada dentro de um contexto social e cultural nunca antes explorado pela literatura.

Para tanto, as sessões de gravações eram divididas em quatro momentos. Em primeiro lugar, antes de iniciar as gravações, um termo de consentimento livre e esclarecido (TCLE), em que era explicada a natureza e as características da pesquisa, era apresentado aos participantes. A esse respeito, é preciso ressaltar que, em São Tomé e Príncipe, não há Comitê de Ética que estabeleça diretrizes para as pesquisas com seres humanos, todavia, as autoridades locais estavam cientes da condução da pesquisa e a Secretária Regional dos Assuntos Sociais e Capital Humano do governo do Príncipe autorizou, oficialmente, o uso dos dados gravados para estudos científicos.

Seguidamente à apresentação do TCLE, a sessão continuava através do levantamento de informações sociolinguísticas sobre os participantes: idade; local de nascimento; no caso de ter nascido em Cabo Verde, a idade de migração para STP; escolaridade e línguas faladas<sup>5</sup>. Isso feito, iniciava-se uma entrevista com narrativas e perguntas sobre temas diversos, como a rotina de trabalho, a migração para STP, a infância, entre outros, com duração média de 30 a 40 minutos. Por fim, a sessão era finalizada com a implementação de um questionário que visava a tradução de determinadas palavras em kabuverdianu, por exemplo, 'mão'. Coletamos o total de 523 palavras que podem ser conferidas no apêndice.

Como especificado anteriormente, o objetivo geral do trabalho de campo foi a coleta de informações estruturais sobre aspectos fonético, fonológico, morfológico, sintático e semântico do KVP que possibilitassem sua descrição linguística inicial. Portanto, os dados não foram coletados com o objetivo específico de um exame fonético como proposto neste artigo. Em decorrência disso, não foram realizados experimentos que permitissem controlar o contexto dos dados elicitados, o que pode causar variabilidade nos valores do primeiro e do segundo formantes das vogais

---

<sup>4</sup> Itens que se referem a conceitos considerados universais, como pronomes pessoais, partes do corpo, cores, denominações para elementos e fenômenos da natureza (como *rio, lagoa, lua, sol, água, terra, pedra, montanha, chuva, calor, frio*), verbos que remetem aos sentidos e às necessidades básicas (como *ver, ouvir, comer, beber, dormir*), nomes como *homem, mulher, criança, pessoa, pássaro*.

<sup>5</sup> Com relação aos aspectos linguísticos, as perguntas versaram sobre a primeira língua adquirida na infância, a língua mais falada, quando e como se deu o aprendizado do português e os contextos de uso de cada uma das línguas

examinadas neste artigo (cf. Barbosa; Madureira, 2015). Dentre os itens investigados, as vogais-alvos podiam ser antecedidas por diferentes qualidades consonantais, bem como pertencerem a diferentes estruturas silábicas. Essas informações assim como o número de dados examinados são especificados na tabela 1<sup>6</sup>.

Tabela 1 - Número de ocorrências investigadas e contextos silábicos e segmentais de acordo com as vogais tônicas analisadas do KVP.

<i>Vogal</i>	<i>N. de ocorrências / porcentagem</i>	<i>Estruturas silábicas</i>	<i>CF1<sup>7</sup></i>	<i>CF2</i>
[i]	69 / 13,2%	CV, CVC, V, CCV, VC	[b t d] [m n] [f v] [l r] [pr br]	[s] (coda)  sílabas seguintes: [b d k g] [m] [v] [tʃ dʒ] [r] [a]
[e]	72 / 13,8%	CCV, CV, CVC, V, CCCV, CCVC	[p b t d k] [m] [v s ʒ] [tʃ] [r] [pr tr] [sk str sfr skr]	[r] (coda)  sílabas seguintes: [p b t d k g] [s ʃ] [l r] [a]
[ɛ]	74 / 14,1%	CV, CVC, CCV, CCCV	[p b t k] [m n] [s] [dʒ] [r] [pr kr] [sfr]	[r s] (coda)  sílabas seguintes: [b t d k g] [m n] [s] [dʒ] [l r] [dr]
[ə]	21 / 4%	CV, CVC, CCV, CVG	[p d k] [m, n] [f s ʒ] [l r] [fl br] [st]	[s ʃ r j w] (coda)  sílabas seguintes: [t] [s ʒ] [tʃ] [r]
[ɐ]	37 / 7,1%	CCVC, CVC, CV, CVG, CGV, CCV, V	[p d k] [m] [f s] [l r] [fl tr] [kw]	[s ʃ r j w] (coda)  sílabas seguintes: [p b t d k] [n] [z] [l r]
[a]	108 / 20,7%	CVC, CCVC, CV, CCV, CVG, CGV, VC	[p b t d k g] [m n] [f s z]	[ʃ r j] (coda)  sílabas seguintes:

<sup>6</sup> As vogais da tabela 1 foram especificadas a partir de análise acústica, como será evidenciado na seção 4.

<sup>7</sup> CF1 = Contexto fonético antecedente; CF2 = Contexto fonético seguinte.

			[tʃ dʒ] [l r] [br tr fl fr] [kw]	[p b t d k g] [n] [v s z ʒ] [l ʎ r] [gr]
[ɔ]	43 / 8,2%	CV, CVC, V	[p b d k] [m n ɲ] [f s] [l r]	[s ʃ r] (coda)  sílabas seguintes: [b t d k] [m] [tʃ dʒ] [r] [br]
[o]	66 / 12,6%	CV, CVC, V, CCVC, CCV	[p d k g] [n] [f s] [tʃ] [r] [fl]	[s ʃ r] (coda)  sílabas seguintes: [b t d k] [m] [tʃ dʒ] [r] [br]
[u]	33 / 6,3%	CVC, CV, CCV	[p t d k] [n] [s ʃ ʒ] [tʃ] [l] [fr]	[s ʃ r] (coda)  sílabas seguintes: [p b t d g] [m] [v s ʒ] [r] [a]

Fonte: Elaboração própria.

No que diz respeito à percepção dos segmentos vocálicos, deve-se destacar que o português, L1 das pesquisadoras, certamente influencia na percepção dos segmentos do kabuverdiano. O fato de o português apresentar duas vogais médias-baixas e duas médias-altas, por exemplo, pode levar a uma identificação dessas mesmas vogais no kabuverdiano. Do mesmo modo, a ausência de [ə] e [ɐ] em posição tônica pode dificultar a discriminação dessas vogais no kabuverdiano. Assim, considerando tais fatores, a definição das vogais levou em conta, além da oitiva, os valores dos formantes (que serão discutidos na subseção 4.1), os quais foram de suma importância inclusive para dirimir dúvidas em alguns contextos.

A partir da tabela 1 vemos que o número de dados, para cada qualidade vocálica, por não ser proveniente de um experimento destinado ao exame acústico dos vocóides em KVP, não está equilibrado. Em decorrência da alta variabilidade de contextos fonético-fonológicos e, mesmo, da distribuição dos dados, os resultados contemplados neste estudo podem estar superestimados ou subestimados no que diz respeito, especificamente, aos valores médios aqui apresentados.

Ainda a esse respeito, é preciso ressaltar que o trabalho de campo nas roças<sup>8</sup> do Príncipe trazia alguns desafios acústicos, na medida em que algumas gravações

<sup>8</sup> Conforme mencionado, as roças constituem antigas propriedades monocultoras, cujas casas grandes foram convertidas em hotéis de luxo. Em muitos casos, os antigos alojamentos dos contratados (as chamadas senzalas) mantiveram-se como a moradia dos caboverdianos e seus descendentes, sendo

foram realizadas em ambiente aberto e, ainda que houvesse um cuidado com barulhos externos, a natureza da gravação, por vezes, impactou a qualidade dos dados. Diante dessa questão, quando o ruído ao fundo era muito intenso e envolvia outras pessoas conversando, animais e barulhos de carro, o token sob análise era excluído visto que a análise acústica estava prejudicada.

Para mitigar os efeitos do desbalanceamento do corpus, apresentamos o desvio padrão dos valores, sendo ideal que, no futuro, novos dados sejam coletados e analisados, tendo em vista o propósito específico de análise acústica. Por ora, assumindo o caráter exploratório e inovador deste estudo, consideramos que, apesar das limitações aqui especificadas, o corpus coletado e a análise empreendida tornam-se relevantes dentro dos estudos fonéticos e fonológicos que visem analisar o kabuverdianu do Príncipe, sendo um passo inicial para trabalhos futuros.

### 3.2 Participantes

Os quatro participantes da pesquisa eram homens adultos, com idades entre 48 e 62 anos, e bilíngues em português-kabuverdianu. Todos eles afirmaram ter aprendido o kabuverdianu na infância, sendo a primeira língua aprendida, tendo contato com o português somente na chegada à escola. Eles mencionaram também se sentirem mais confortáveis em usar o kabuverdianu do que o português, geralmente recorrendo ao português para comunicação com pessoas que não fazem parte da comunidade caboverdiana. Além disso, os quatro eram filhos de caboverdianos que migraram para o Príncipe, sendo que dois participantes nasceram em Cabo Verde e os outros dois nasceram no Príncipe. O perfil sociolinguístico dos participantes pode ser observado no quadro 1:

Quadro 1 - Perfil sociolinguístico dos entrevistados.

Participante	Origem	Idade em 2018	Idade de migração para o Príncipe	Tempo no Príncipe	Moradia
01	Cabo Verde	62	8	54 anos	Sundy
02	Cabo Verde	51	2	49 anos	Belo Monte
03	Príncipe	43	---	---	Belo Monte
04	Príncipe	48	---	---	Belo Monte

Fonte: Elaboração própria.

A partir do quadro 1, observam-se algumas diferenças sociolinguísticas entre os participantes da pesquisa, que podem impactar nos resultados. Por exemplo, dois participantes nasceram em Cabo Verde e dois já nasceram no Príncipe, cenário que implica idades diferentes quanto à aquisição da linguagem, embora todos os falantes estivessem dentro do período crítico ao adquirir o kabuverdianu com alto grau de fluência (cf. Gleitman; Newport, 1995). Ademais, só foi possível coletar dados com homens, já que as mulheres não estavam disponíveis para participar da pesquisa na curta estadia no Príncipe. Outro fator a ser considerado é a moradia dos participantes. Três deles vivem na roça Belo Monte, enquanto um deles vive em Sundy, não tendo sido possível coletar mais dados com falantes de uma mesma comunidade. Apesar de reconhecermos a relevância dessas questões sociais, não consideraremos nesta análise

---

residências simples, algumas vezes compostas de poucos cômodos (sala e quarto(s)), com a cozinha e o banheiro em uma área externa.

a interferência de tais variáveis dada a natureza preliminar do corpus coletado, deixando-se para um estudo futuro esse exame.

### 3.3 Procedimento de Análise

Para descrição e análise dos aspectos acústicos constituintes das vogais tônicas do KVP, os seguintes critérios foram considerados: (i) a duração em milissegundos; (ii) o primeiro formante (F1), relacionado à altura da língua na produção vocálica; (iii) o segundo formante (F2), correspondente ao movimento horizontal da língua, isto é, a posterioridade/anterioridade da língua.

A duração foi mensurada considerando o aparecimento da energia de formante mais intensa (limite esquerdo) e sua atenuação (limite direito), particularmente da energia de F2 (Barbosa; Madureira, 2015), ao passo que os valores de F1 e F2, em Hertz, foram extraídos a partir de porções mediais das vogais, posição em que não há a interferência de sons circunvizinhos e de pausa. Essa análise acústica das vogais-alvo foi realizada através do Praat (Boersma; Weenick, 2022) e, posteriormente, os valores obtidos foram normalizados através do z-score para a duração e do método Lobanov para os formantes.

Em nossa análise, a comparação do KVP foi feita com o kabuverdiano de Santiago – com base em estudos como Quint (2000), Lang (2001, 2002) e Rodrigues (2007) – pelo fato de essa ser a procedência da maior parte dos caboverdianos que chegaram à ilha do Príncipe segundo registros como Carreira (1983), como apontado na seção 2. Em estudos posteriores, a comparação poderá ser ampliada para as variedades de outras ilhas.

Isso posto, apresentamos, na seção 4, a seguir, nossos resultados.

## 4 VOGAIS TÔNICAS ORAIS DO KVP

Em posição tônica, no kabuverdiano do Príncipe, foram encontrados nove fones vocálicos orais ([i, e, ɛ, ə, ɐ, a, ɔ, o, u]), estabelecidos com base nos formantes (que serão analisados na subseção seguinte), como se pode observar nos exemplos no quadro 2:

Quadro 2 - Vogais Tônicas do kabuverdiano do Príncipe.

<i>Vogal</i>	<i>Item</i>	<i>Transcrição</i>	<i>Glosa</i>
[i]	vivi <sup>9</sup>	[ˈvi.vi]	‘viver’
[e]	direta	[diˈre.tə]	‘direita’
[ɛ]	brumédju	[bru.ˈmɛ.dʒo]	‘vermelho’
[ə]	nasi	[ˈnə.sɪ]	‘nascer’
[ɐ]	kaza	[ˈkɐ.zə]	‘casa’
[a]	brasu	[ˈbra.su]	‘braço’
[ɔ]	róstu	[ˈrɔs.tu]	‘rosto, face’
[o]	korta	[ˈkof.tə]	‘cortar’
[u]	juga	[ˈʒu.gɐ]	‘jogar’

Fonte: Elaboração própria.

<sup>9</sup> A grafia dos dados do KVP segue o ALUPEC, proposta de um alfabeto unificado para as ilhas de Cabo Verde como um primeiro passo para a padronização da língua (Cabo Verde, 1998).

Esse inventário fonético é, de forma geral, semelhante àquele vigente para a variedade de Santiago. De acordo com Quint (2000) e Lang (2001, 2002), o kabuverdianu de Santiago teria oito fonemas vocálicos: /i, e, ε, ɐ, a, ɔ, o, u/. Rodrigues (2007) também aponta oito segmentos, contudo em lugar de /ɐ/, a autora menciona o /ə/. Para defender a oposição entre as vogais centrais (terminologia empregada pelos autores e que adotaremos aqui), Quint (2000) e Lang (2001, 2002) se apoiam na existência de pares mínimos como *sabi* ['sɐ.bi] 'saber' e *sábi* ['sa.bi] 'agradável'. No que tange às realizações fonéticas, os três autores apontam oito realizações possíveis: [i, e, ε, ə/ɐ, a, ɔ, o, u]. Comparando as duas variedades, observa-se que o kabuverdianu do Príncipe amplia o leque de vogais centrais possíveis em posição tônica, sendo possíveis três realizações: [ə, ɐ, a] (que são usadas pelos quatro falantes desta pesquisa), enquanto que, em Santiago, são mencionadas apenas duas ([a] e [ɐ] ou [a] e [ə])<sup>10</sup>.

Quadro 3 - Vogais Tônicas Fonéticas no kabuverdianu de Santiago e da Ilha do Príncipe.

Kabuverdianu de Santiago			Kabuverdianu do Príncipe		
Quint (2000) e Lang (2001, 2002)		Rodrigues (2007)	Autoras (análise acústica)		
i	_____ u	i _____ u	i	_____ u	
e	_____ o	e _____ o	e	_____ o	
ε	_____ ɔ	ε [ə] ɔ	ε	[ə] ɔ	
	[ɐ] a	a		[ɐ] a	

Fonte: Elaboração própria.

Do ponto vista acústico, foram analisados os seguintes correlatos das vogais fonéticas orais do kabuverdianu do Príncipe: o primeiro e o segundo formantes (F1 e F2) e a duração, cujos resultados serão apresentados nas subseções seguintes.

#### 4.1 Análise do primeiro e do segundo formantes

Os valores normalizados de F1 e F2, em Hertz, das vogais tônicas do kabuverdianu do Príncipe são apresentados na tabela 2:

Tabela 2 - Valores Médios Normalizados de F1 e F2 (Hz) das vogais tônicas do kabuverdianu do Príncipe. DV: desvio padrão.

Vogal	F1	DV	F2	DV
[i]	346	31,3	1974	146,1
[e]	411	34,1	1837	121,7
[ε]	486	51,2	1781	126,3
[ə]	480	42,3	1463	117,1
[ɐ]	551	39,8	1405	167
[a]	626	36,2	1392	213
[ɔ]	520	43,3	1095	96,9
[o]	422	36,2	1086	140,3
[u]	378	26,6	1194	217,3

Fonte: Elaboração própria.

Com relação às diferenças entre F1 e F2 para as vogais médias, observa-se que para as anteriores, a diferença entre os valores de F1 é de 75 Hz e entre os de F2 é de

<sup>10</sup> Os estudos citados para o kabuverdianu de Santiago não realizaram análises acústicas.

56 Hz; já para as posteriores (nomenclaturas usadas pelos autores que se debruçaram sobre o kabuverdianu de Santiago e que manteremos neste artigo), a diferença entre os valores de F1 é de 98 Hz e entre os de F2 é de 9 Hz. Os valores de F1 são menores do que os números reportados pela autora para diferentes variedades do português – santomense (PST), principense (PP) e europeia (PE) –, como se observa na tabela 3, que apresenta também os números para o kabuverdianu do Príncipe (KVP).

Tabela 3 - Diferença entre os valores de F1 (Hz) das vogais tônicas médias-altas e médias-baixas no PST, PP, PE e KVP.

Vogal	Variedade							
	PST		PP		PE		KVP	
	F1	Dif.	F1	Dif.	F1	Dif.	F1	Dif.
[e]	386	115	366	138	355	100	411	75
[ɛ]	501		504		455		486	
[o]	413	124	397	165	363	128	422	98
[ɔ]	537		562		491		520	
[e]	386	83	366	113	355	71	411	65
[i]	303		253		284		346	
[o]	413	74	397	98	363	60	422	44
[u]	339		299		303		378	
[ə]	530	174	521	206	—	—	480	ə/v: 71
[ɐ]	—		—		—		551	ə/a: 146
[a]	704		727		661		626	v/a: 75

Fonte: Elaboração própria com base em dados de Balduino (2022) para o PST e o PP e de Escudero et al. (2009) para o PE.

A partir da tabela 3, observa-se que, dentre as línguas analisadas, no kabuverdianu do Príncipe, a diferença entre as médias é menor, o que sugere que, nessa língua, essas vogais estão mais próximas entre si. De forma geral, a diferença entre todas as vogais é menor no kabuverdianu, o que pode se refletir em variação, uma vez que o espaço acústico é menor e, em decorrência disso, as vogais tendem a estar mais próximas, podendo ter valores de F1 sobrepostos, como se verá mais adiante.

Sobre a realização das vogais médias, em muitos casos, houve dúvidas sobre qual vogal era realizada, sendo necessário ouvir mais de uma vez e inclusive consultar outros linguistas para definir a sua pronúncia. Nesses casos, a fim de confirmar ou não a realização obtida através da oitiva, recorreu-se aos valores dos formantes. Além disso, em uma das ocorrências da palavra **berdi** ‘verde’, houve dúvidas se a vogal tônica era [i] ou [e] e os valores dos formantes também não foram suficientes para se chegar a uma definição. Já em uma das realizações da palavra **sol** ‘sol’, a pronúncia da vogal parece começar sendo [ɔ] e depois passar para [o], sugerindo uma gradiência na pronúncia. A esse respeito, vale mencionar que esse movimento gradiente pode ser motivado pelo movimento articulatório para a realização da coda [ʰ] que apresenta um alteamento da mandíbula em relação a [ɔ] e não, necessariamente, por uma gradiência resultante da variação de [ɔ] e [o] em diferentes contextos segmentais – uma hipótese a ser confirmada, futuramente, diante de mais dados. Nesses dois casos, optou-se por excluir os dados da análise.

Acerca dessas possíveis realizações intermediárias, Ladefoged (1982 [1975]) aponta que, entre vogais, é muito mais comum haver tais realizações do que entre consoantes, mencionando a possibilidade de haver uma vogal produzida entre uma vogal alta e uma vogal média. Para ele, os casos em que não há oposição entre as vogais

favoreceriam ainda mais realizações intermediárias. Além disso, as vogais médias não estariam na mesma altura em línguas diferentes, sendo possível encontrar variações e inclusive realizações não propriamente equidistantes. A esse respeito, Quint (2000) e Lang (2002) mencionam que, no kabuverdianu de Santiago, é possível a realização de uma vogal tônica entre as médias-altas e as médias-baixas (tanto para as coronais quanto para as dorsais). Tal afirmação, que pode ser uma explicação válida para o cenário do kabuverdianu, precisa ser melhor analisada em estudos futuros.

Do ponto de vista perceptual, não se pode descartar que o fato de as pesquisadoras serem falantes de português como L1 (língua que possui duas vogais médias-baixas e duas médias-altas) exerce influência na percepção dos segmentos do kabuverdianu. Frota e Vigário (2000), em seu estudo acerca da percepção do acento em português europeu e brasileiro, mostram que os padrões acentuais da língua materna guiam a percepção dos acentos em uma outra variedade. Desse modo, o padrão rítmico do português europeu é projetado por um falante português em sua percepção de dados do português brasileiro e vice-versa. Seguindo essa mesma linha, Mongelo, Seara e Agostinho (2021), ao focar em uma informante falante nativa do guarani que aprende português brasileiro como terceira língua (sendo o espanhol a segunda língua), apontam que, na produção das vogais do português, é reproduzido o sistema de alturas da L1, o guarani, que possui um sistema com 3 contrastes de altura, e não 4, como o português. Assim, além de olhar para aspectos da produção, as dificuldades em determinar se a vogal era [e] ou [ɛ], [o] ou [ɔ], [ə], [ɐ] ou [a], no kabuverdianu do Príncipe devem ser analisadas ainda no que tange à percepção de falantes bilíngues ou monolíngues, aspecto ainda pouco considerado nas análises fonético-fonológicas. Diante desse cenário, aponta-se a necessidade de estudos futuros mais acurados a fim de analisar a realização fonética das vogais médias, que poderá revelar realizações intermediárias, e seus reflexos no sistema fonológico, como será brevemente discutido ainda nesta seção.

O quadro fonético encontrado para o kabuverdianu do Príncipe pode ser relacionado com discussões de variedades do português e outras línguas em virtude de algumas similaridades estruturais e históricas. Quanto às vogais centrais em posição tônica, segundo Balduino (2022), no português do Príncipe, aparece somente [a], não sendo verificado o [ə], posicionamento revisto em estudo posterior (Balduino; Freitas no prelo), que aponta a presença de schwa na fala dos homens do Príncipe. Já no português de São Tomé, tanto [a] quanto [ə] foram encontrados na posição tônica. Em português europeu, há três variantes possíveis: [ə], [ɐ], [a]<sup>11</sup>. Discutindo variedades do português em África, Agostinho (no prelo) menciona que, no português de Moçambique, há diferentes realizações da vogal central em posição tônica; e, no português de Angola, uma realização possível é [ɐ] em posição tônica. Considerando esse cenário e o fato de que as línguas crioulas costumam ter inventários fonológicos com menos segmentos do que suas línguas lexicadoras (Velupillai, 2015), é possível conjecturar que as diferentes realizações das vogais centrais no kabuverdianu do Príncipe sejam alofones, contudo somente um estudo posterior poderá confirmar ou refutar tal hipótese.

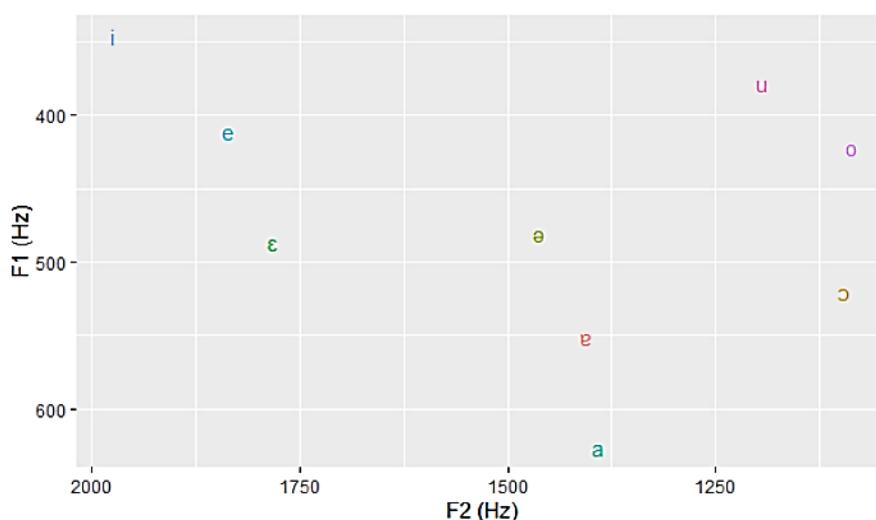
Passando para a análise das vogais médias coronais e dorsais, Wetzels (1992) menciona que, no português brasileiro, a carga funcional das oposições entre [e] e [ɛ],

---

<sup>11</sup> A relação do KVP com o PST, o PP e o PE foi discutida na seção 2.

[o] e [ɔ] seria baixa, dada a sua ocorrência somente em sílabas acentuadas e em um número pequeno de pares mínimos. Esse mesmo cenário é apontado por Hall (2013) para as vogais médias coronais do francês e do italiano. Citando o trabalho de Ladd (2006), o autor afirma que o contraste entre [e] e [ɛ] existe somente em posição tônica, sendo neutralizado nas demais posições, em que é possível uma realização intermediária, entre os dois segmentos. Além disso, Balduino (2022) e Agostinho (no prelo) apontam uma perda de contraste entre as médias em posição tônica em variedades africanas do português, como no português de São Tomé e Príncipe (cérebro [ˈsɛ.ri.bʊ] ~ [ˈse.ri.bʊ] e abóbora [a.ˈbɔ.burɐ] ~ [a.ˈbo.burɐ]), no português de Angola (avô e avô [a.ˈvɔ]). Considerando esse cenário, Agostinho (no prelo) sugere a possibilidade da fusão das vogais médias, que assim deixariam de opor significado. Essa afirmação se apoia em trabalhos como Wendel, Kaplan e Jackson (2013) e Hualde (2014), autores que defendem a chamada hipótese da carga funcional: pares de fonemas com baixa carga funcional (isto é, responsáveis por distinguir um número reduzido de vocábulos) estariam mais propensos de sofrer uma mudança, como neutralizações. Do ponto de vista da percepção, Demolin (2012) aponta que os falantes têm a tendência de evitar contrastes perceptualmente fracos, o que pode ser mais um aspecto favorecendo a fusão entre fonemas. Essa baixa carga funcional da oposição entre as vogais médias ilustra que as fronteiras entre fonema e alofone não são tão nítidas quanto se pode pensar à primeira vista. A esse respeito, Goldsmith (1995) aponta que apenas os rótulos de fonema e fone não conseguem explicar a realidade encontrada nas línguas, sendo necessário postular um continuum entre essas duas categorias, com segmentos que ora se aproximam mais dos fonemas, ora dos fones. Hall (2013) discute esse cenário (e inclusive o continuum proposto por Goldsmith) e menciona que os critérios comumente usados para distinguir fonema de fone – como previsibilidade da distribuição, distinção lexical ou teste da comutação, julgamento do falante nativo, alternâncias, similaridade fonética, ortografia, lugar no sistema/comportamento fonológico – possuem problemas não só internos, como na sua relação uns com os outros. O autor conclui que já se reconhece que há categorias entre fonemas e fones (conhecidas na literatura por termos diversos como quasi-fonemas/quasi-contrastes, semifonemas/semicontrastes, contrastes fracos/parciais, entre outros), contudo ainda não se chegou a uma explicação suficientemente clara sobre esses elementos intermediários (Hall, 2013). Essa discussão poderá ser aplicada em estudos futuros para definir o estatuto das vogais médias no kabuverdianu do Príncipe, estabelecendo se existe de fato oposição e qual sua carga funcional.

A partir das médias de F1 e F2, foi elaborado o gráfico da dispersão acústica das vogais orais em posição tônica, conforme figura 1:



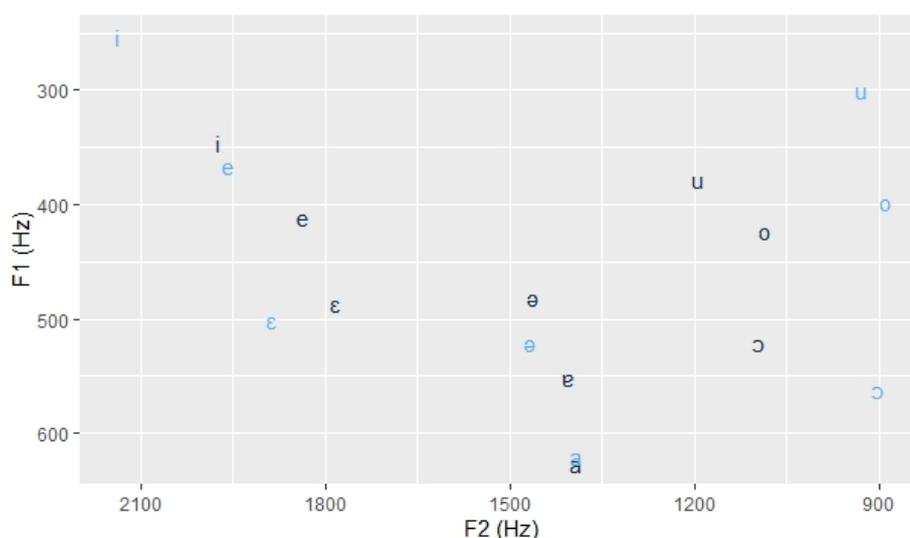
Fonte: Elaboração própria.

Figura 1 - Gráfico do espaço acústico (F1 x F2) dos valores médios de vogais orais tônicas do kabuverdiano do Príncipe.

Quanto à análise das vogais, Ladefoged (1982 [1975]) chama atenção para o fato de que esta deve ser feita de forma relativa, sempre considerando o comportamento e a posição de uma vogal com relação à outra. Ademais, considerando o espaço acústico, o autor menciona que “[...] as vogais próximas das extremidades do gráfico são mais distintas umas das outras do que as vogais do meio, e as diferenças na qualidade das vogais tornam-se progressivamente reduzidas entre as vogais mais próximas do centro.” (Ladefoged, 1982 [1975], p. 79; tradução nossa). Essa assunção do autor parece ser corroborada pelos dados do kabuverdiano do Príncipe, na medida em que as vogais [e] e [ɛ], [o] e [ɔ], [a], [ɐ] e [ə] tendem a ser mais próximas entre si, guardando similaridades<sup>12</sup> (o que se verifica, inclusive, pela dificuldade, em alguns casos, em determinar qual vogal foi realizada exatamente – como discutido na tabela 3, as frequências entre essas vogais são mais próximas que em variedades do português), do que as vogais altas [i] e [u] com relação ao [a].

Os valores médios de F1 e F2 obtidos para o kabuverdiano do Príncipe foram comparados com aqueles encontrados para o português do Príncipe (dados retirados de Balduino, 2022), língua também falada amplamente pelos informantes desta pesquisa. A comparação entre os valores pode ser vista na figura 2, em que o espaço acústico do kabuverdiano encontra-se em vermelho e o do português em azul (dados coletados com falantes diferentes):

<sup>12</sup> Um desdobramento futuro para ampliação do estudo seria comparar as realizações vocálicas do kabuverdiano e do português dos mesmos falantes do Príncipe, o que poderia esclarecer alguns aspectos ainda não compreendidos do KVP.



Fonte: Elaboração própria.

Figura 2 - Gráfico do espaço acústico (F1 x F2) dos valores médios de vogais orais tônicas do kabuverdianu (em azul escuro) e do português (em azul claro) do Príncipe.

Observa-se que enquanto o kabuverdianu possui 9 segmentos, o português possui somente 7, sendo a diferença relacionada às vogais centrais. Balduino (2022) e Santiago et al. (2022) apontam que essa variedade do português possui sete fonemas vocálicos orais em posição tônica (/i, e, ɛ, a, ɔ, o, u/), sendo possível ocorrer mesmo nessa posição neutralização da oposição entre as vogais médias altas e baixas. Os espaços vocálicos acústicos têm sido analisados, entre outros aportes teóricos, a partir da Teoria da Dispersão, proposta por Flemming (2004), segundo a qual os contrastes fonológicos buscam atender três princípios: (i) maximizar contrastes; (ii) minimizar o esforço articulatório e (iii) maximizar a distintividade dos contrastes. Ainda que, neste trabalho, não tenha sido possível propor o quadro vocálico fonológico do kabuverdianu do Príncipe, pode-se aplicar alguns aspectos da teoria às realizações fonéticas vocálicas da língua, bem como podem dar ensejo a pesquisas futuras sobre o tema. Nota-se, por exemplo, que, mesmo tendo mais segmentos, o espaço acústico do kabuverdianu do Príncipe é mais comprimido do que aquele do português, que possui menos segmentos. O fato de o kabuverdianu possuir tal configuração pode ser um dos motivos que leve a confusões entre alguns fones (especialmente na série das vogais médias) para não falantes da língua, sobretudo, no caso das pesquisadoras falantes de português, idioma que distingue [e] x [ɛ] e [o] x [ɔ] em sílaba tônica, já que, de acordo com Kingston (2007, p. 412; tradução nossa), “[...] se um inventário for muito grande ou seus membros estiverem acusticamente muito próximos uns dos outros, mensagens distintas serão confundidas umas com as outras.” Essa é uma hipótese a ser testada em estudos futuros em relação aos falantes do KVP.

Um resultado diferente é trazido por Marusso (2016) a partir da comparação dos espaços acústicos do português (que possui 7 vogais orais) e do inglês (que apresenta 11 vogais orais). Os resultados encontrados contrariam o previsto pela teoria na medida em que mostram que o inglês, cujo inventário é maior, possui realizações menos precisas e mais variáveis que o português. Além disso, o espaço acústico do português é maior do que o do inglês, com uma maior dispersão das vogais. Essa breve discussão revela que a dispersão é um aspecto que precisa ser retomado em pesquisas

futuras, inclusive com dados de fala controlada, a fim de obter uma análise mais acurada do espaço acústico do kabuverdiano do Príncipe.

No que diz respeito ao comportamento das vogais centrais e médias, em diversas variedades de kabuverdiano, existem discussões quanto a suas realizações e seu estatuto fonológico. Quint (2000) aponta que as vogais [e], [ɛ], [o] e [ɔ] da variedade de Santiago se realizam de forma mais baixa (aberta) do que em português. Ademais, segundo o autor, em posição final absoluta ou quando há uma consoante ([ʃ, w, j, r, l, s/]) na posição de coda da sílaba, ocorreria neutralização das vogais médias e centrais da língua, com um sistema com 5 fonemas (com 8 alofones: [i, e, ɛ, ɐ, a, ɔ, o, u]). Rodrigues (2007) aponta uma diferença dialetal, com as variedades de Santiago e Fogo dando preferência às realizações médias-baixas, enquanto que, em Santo Antão e São Vicente, predominam as realizações médias-altas (conforme mencionado na seção da metodologia, o foco nesse artigo recaiu na variedade de Santiago). Quanto à variedade falada na ilha do Maio, Moreira (2014) menciona que só foi encontrado um par mínimo entre /a/ e /ɐ/ ([ˈkɛl] ‘cal’ e [ˈkal] ‘qual’); nos casos das vogais médias, nos pares analisados, além da mudança da vogal, houve também alteração no acento (paroxítono e oxítono), o que faz com que a diferença esteja em dois aspectos e não em apenas um, tratando-se de pares análogos (como [re.ˈgɐ] ‘regar’ e [ˈrɛ.gɐ] ‘rega’ e [ko.ˈbɐ] ‘cavar’ e [ˈkɔ.bɐ] ‘cova’). Já no que tange ao kabuverdiano de Santo Antão (Baptista, 2014) e do Fogo (Moreira, 2020), foram registrados pares mínimos para as vogais médias anteriores e posteriores e as vogais centrais.

Considerando especificamente a variedade de Santiago, Rodrigues (2007) apresenta pares mínimos com as médias-altas e médias-baixas anteriores e posteriores que justificariam a existência de quatro vogais médias como fonemas em kabuverdiano, como se observa em (1) (último exemplo retirado de Lang, 2002):

- (1) a. *séra* [ˈsɛ.rɐ] ‘1. serra (ferramenta), 2. serra (montanha), 3. cera’ e *sera* [ˈse.rɐ] ‘serrar’  
 b. *féra* [ˈfɛ.rɐ] ‘animal selvagem, fera’ e *fera* [ˈfe.rɐ] ‘1. festa popular, 2. mercado, feira’  
 c. *óra* [ˈɔ.rɐ] ‘hora’ e *ora* [ˈo.rɐ] ‘orar, rezar’  
 d. *róda* [ˈrɔ.dɐ] ‘1. roda, 2. (grande) quantidade, 3. arredores’ e *roda* [ˈro.dɐ] ‘1. rodar, girar, 2. rodear, circundar’

Quanto às realizações das vogais médias-baixas, segundo Rodrigues (2007), tais vogais geralmente ocorrem na sílaba tônica de palavras paroxítonas que terminam em vogal central, como em *kabésa* [kɐ.ˈbɛ.sɐ] ‘cabeça’, *manéra* [mɐ.ˈnɛ.rɐ] ‘maneira, modo’, *góta* [ˈgɔ.tɐ] ‘gota’, *bóka* [ˈbɔ.kɐ] ‘boca’ (exemplos retirados da autora). A autora menciona ainda que, nas variedades de Sotavento, há alguns poucos casos de vogais médias-baixas em palavras que terminam em /u/ ou /r/: *sédu* [ˈsɛ.du] ‘cedo, logo’, *xofér* [ʃo.ˈfɛf] ‘motorista’ (exemplos retirados da autora e de Lang, 2002).

Passando para as vogais centrais, autores como Quint (2000, 2001), Lang (2002) e Rodrigues (2007) consideram a existência de dois fonemas na língua. De acordo com Quint (2000), apesar de não ser muito recorrente, a oposição entre as vogais centrais permite distinguir pares mínimos frequentes, como os que constam no exemplo (2). Para ele, essa diferença entre /a/ e /ɐ/ é exclusiva do kabuverdiano falado em Santiago, não aparecendo em outras variedades. Quint (2001) advoga que a

incorporação do segmento /ɐ/ ao quadro vocálico serviria para marcar a distinção entre nomes e verbos. Assim sendo, haveria uma motivação morfológica para a inserção da vogal /ɐ/ no sistema. O autor menciona os pares mínimos em (2) envolvendo as vogais centrais:

(2) a. parti ['pɛr.ti] 'partir (1. quebrar, 2. repartir, 3. ir embora)' e parti ['par.ti] 'parte'

b. sabe ['sɛ.bi] 'saber' e sabi ['sa.bi] '1. alegria, 2. agradável'

Já Rodrigues (2007) afirma que, na variedade de Santiago, em posição tônica, o [a] é mais comum, ao passo que, em posição átona, é mais frequente o [ə]. Para ela, em posição tônica, /a/ e /ə/ devem ser considerados fonemas, sendo a oposição neutralizada em posição átona (alofones posicionais). No que diz respeito ao uso de cada um dos segmentos centrais, Lang (2002) menciona que há alguns contextos em que [a] é a forma preferida e outros em que a preferência incide sobre o [ɐ]. Assim, o [a] é mais escolhido nos seguintes contextos: (i) ditongo [aʊ] em posição final de palavra: *máu* ['maʊ] 'mau', *kularáu* [ku.lɛ.'raʊ] 'colorau' (exemplo retirado de Lang, 2001); (ii) sílabas tônicas em posição final terminadas em [s]: *aliás* [ɛ.'ljas] 'aliás'. Já o [ɐ] é mais usado nos seguintes casos: (i) ditongo [ɛɪ] em final de palavra: *papai* [pɛ.'paɪ] 'papai'; (ii) sílabas tônicas em posição final terminadas em [t]: *sal* ['sɛt] 'sal'; (iii) monossílabos que têm um onset complexo cujo elemento final é uma líquida: *fla* ['flɛ] 'falar', *tra* ['trɛ] 'tirar, extrair' (exemplos retirados desse autor).

Ainda com relação à altura das vogais, Lang (2001, 2002) e Rodrigues (2007) mencionam que um aspecto particular do kabuverdiano consiste em permitir identificar a que classe gramatical uma determinada palavra (que tem na língua uma outra forma quase homófona) pertence a partir da altura de sua vogal tônica. Em geral, os itens verbais teriam como vogal tônica uma vogal alta ou central/média-alta, enquanto que os itens nominais teriam uma vogal alta ou central/média-baixa. Há inclusive casos de pares mínimos cuja oposição se dá somente com relação à altura de suas vogais tônicas. É o que se vê, por exemplo, em (3) (exemplos retirados de Lang, 2002):

(3) a. *careka* [kɛ.'re.kɛ] 'ficar careca' e *karéka* [kɛ.'rɛ.kɛ] 'careca'

b. *regra* ['re.grɛ] 'pôr em ordem, regularizar' e *régra* ['rɛ.grɛ] 'regra'

c. *karapati* [kɛ.rɛ.'pɛ.ti] 'segurar(-se)' e *karapáti* [kɛ.rɛ.'pa.ti] 'carrapato'

d. *astia* ['ɛs.tjɛ] 'hastear (bandeira)' e *astia* ['as.tjɛ] 'bastão, vara'

e. *diróta* [di.'ro.tɛ] 'derrotar' e *diróta* [di.'rɔ.tɛ] 'derrota'

f. *mostra* ['mos.trɛ] 'mostrar(-se)' e *móstra* ['mɔs.trɛ] 'amostra, prova'

Com relação a essas vogais, conforme mencionado, não foi possível estabelecer seu estatuto fonológico (no caso das vogais centrais, a análise sugere que se tratam de alofones). Por outro lado, buscamos observar a sua distribuição no corpus analisado e se havia algum contexto linguístico inibidor ou favorecedor de cada uma das realizações. No que diz respeito às vogais centrais, foram coletados 166 dados, distribuídos da seguinte forma:

Tabela 4 - Distribuição das vogais centrais.

<i>Vogal</i>	<i>Número de dados</i>	<i>Porcentagem</i>
[ə]	21	13%
[ɐ]	37	22%
[a]	108	65%
TOTAL	166	100%

Fonte: Elaboração própria.

É possível observar que, apesar de a vogal baixa [a] constituir a maioria das realizações da vogal central, foram também encontradas as variantes mais altas, que juntas somam 35% e em geral não costumam ser muito comuns em posição tônica.

Quanto às vogais médias anteriores, foram encontrados 146 dados, divididos quase igualmente:

Tabela 5 - Distribuição das vogais médias anteriores.

<i>Vogal</i>	<i>Número de dados</i>	<i>Porcentagem</i>
[ɛ]	74	51%
[e]	72	49%
TOTAL	146	100%

Fonte: Elaboração própria.

Por fim, com relação às vogais médias dorsais, a preferência foi para a média-alta:

Tabela 6 - Distribuição das vogais médias posteriores.

<i>Vogal</i>	<i>Número de dados</i>	<i>Porcentagem</i>
[ɔ]	43	39%
[o]	66	61%
TOTAL	109	100%

Fonte: Elaboração própria.

Quanto à distribuição, não foi possível definir o que condiciona as diversas realizações, visto que, conforme mencionado na seção 3, a amostra não estava equilibrada. Foram encontrados casos de variação nas realizações das vogais centrais, médias anteriores e posteriores, sendo mais comum em vogais centrais. Registrou-se variação na realização de um mesmo item lexical na fala de um mesmo entrevistado e também na fala de mais de um entrevistado, como se observa no quadro 4 (há outros exemplos de variação das vogais centrais, já com relação às vogais anteriores e posteriores, foram encontrados somente esses exemplos):

Quadro 4 - Exemplos de variação entre as vogais centrais e médias.

<i>Item</i>	<i>Realizações</i>
<b>fla</b> ‘falar’	Entrevistado 1: fl[ə] Entrevistado 3: fl[a] ~ fl[ɐ] Entrevistado 4: fl[a] ~ fl[ɐ] ~ fl[ə]
<b>dimas</b> ‘demais’	Entrevistado 1: dim[ə]s ~ dim[ɐ]s
<b>karni</b> ‘carne’	Entrevistado 3: k[a]rni ~ k[ɐ]rni Entrevistado 4: k[ɐ]rni
<b>kaza</b> ‘casa’	Entrevistado 1: k[a]sa ~ k[ɐ]sa Entrevistado 2: k[a]sa
<b>dá</b> ‘dar’	Entrevistado 2: d[ə] Entrevistado 3: d[a] ~ d[ɐ]

<b>pai</b> ‘pai’	Entrevistado 1: p[ɐ]i Entrevistado 3: p[ə]i ~ p[a]i
<b>bebi</b> ‘beber’	Entrevistado 1: b[ɛ]bi Entrevistados 2 e 3: b[e]bi
<b>kabesa</b> ‘cabeça’	Entrevistado 1: kab[e]sa Entrevistado 2: kab[ɛ]sa
<b>medu</b> ‘medo’	Entrevistado 1: m[e]du Entrevistado 2: m[ɛ]du ~ m[e]du
<b>pretu</b> ‘preto’	Entrevistados 1 e 3: pr[e]tu Entrevistado 2: pr[ɛ]tu
<b>sfrega</b> ‘esfregar’	Entrevistado 3: sfr[e]ga Entrevistado 4: sfr[ɛ]ga
<b>dor</b> ‘dor’	Entrevistado 1: d[o]r Entrevistado 4: d[ɔ]r ~ d[o]r
<b>korta</b> ‘cortar’	Entrevistado 1: k[ɔ]rta ~ k[o]rta Entrevistado 2: k[o]rta
<b>obu</b> ‘ovo’	Entrevistado 1: [ɔ]bu Entrevistado 2: [o]bu
<b>rostu</b> ‘rosto’	Entrevistado 1: r[ɔ]stu Entrevistado 2: r[o]stu

Fonte: Elaboração própria.

Focando em critérios estruturais, observa-se que a série das vogais centrais e médias anteriores e posteriores, além de estar presente em palavras monossílabas e dissílabas ou polissílabas (conforme dados do quadro anterior):

Quadro 5 - Contextos em que aparecem as vogais centrais e médias.

<i>Contexto</i>	<i>Exemplo</i>
Coocorre nos mesmos contextos segmentais antecedentes (CF1)	[kə.] / [kɐ.] / [ka.] [be.] / [be.] [pɔ.] / [po.]
Coocorre nos mesmos contextos segmentais seguintes (CF2)	[ə.r] / [ɐ.r] / [a.r] [e.s] / [e.s] [ɔ.k] / [o.k]
Coocorre em sílabas preenchidas por coda	[səʔ.] / [mɐʔ.] / [saʔ.] / [maʔ.] [kɛr.] / [ker.] [kɔs.] / [kos.]
Coocorre na posição inicial absoluta	#[ɐ.] / #[ə.] #[ɔ.] / #[o.]

Fonte: Elaboração própria.

Dito de outra forma, não houve um contexto segmental ou silábico que favorecesse uma ou outra realização ([e] ou [ɛ], [o] ou [ɔ], [a], [ə] ou [ɐ]). Assim, as distribuições apontadas por Lang (2002) para as vogais centrais – com [a] sendo preferida no caso do ditongo [aʊ] e em sílabas tônicas finais terminadas em [s] e o [ɐ] sendo preferido no ditongo [ɐi] e em sílabas tônicas finais terminadas em [ʔ] – e Rodrigues (2007) para as vogais médias – com as vogais médias-baixas sendo preferidas em palavras terminadas com vogal central – não se fizeram presentes no corpus do kabuverdiano do Príncipe analisado.

Tendo apresentado os resultados para os formantes, a subseção seguinte traz algumas considerações gerais sobre a duração.

## 4.2 Considerações gerais sobre a duração

Os valores normalizados de duração, em milissegundos, das vogais tônicas do kabuverdiano do Príncipe são apresentados na tabela 7:

Tabela 7 - Valores Médios Normalizados da duração (ms) das vogais tônicas do kabuverdiano do Príncipe. DV: desvio padrão.

<i>Vogal</i>	<i>Duração</i>	<i>DV</i>
[i]	98	34,6
[e]	104	26,5
[ɛ]	113	22,6
[ə]	99	43,9
[ɐ]	100	32,6
[a]	119	31,2
[ɔ]	124	28,7
[o]	109	31,1
[u]	86	28,3

Fonte: Elaboração própria.

Considerando os valores da tabela 7, no que diz respeito à duração das vogais, [ɛ, a, ɔ] são mais ligeiramente mais longas quando comparadas às vogais altas [i, u] e mesmo às vogais [e] e [o] (com a diferença entre [ɛ] e [i] de 15 ms; e entre [ɔ] e [u] de 34 ms) (ainda que a diferença aqui seja menor), como tem sido reportado para as línguas em geral (Cristófaró Silva et al., 2019) e para variedades do português (Balduino, 2022). A comparação com as diferentes variedades do português é reportada na tabela 8:

Tabela 8 - Comparação entre os Valores Médios Normalizados da duração (ms) das vogais tônicas do PST, PP, PE e KVP.

<i>Vogal</i>	<i>Variedade</i>			
	<i>PST</i>	<i>PP</i>	<i>PE</i>	<i>KVP</i>
[i]	77	78	84	98
[e]	90	88	97	104
[ɛ]	93	109	106	113
[a]	102	106	108	119
[ɔ]	106	107	104	124
[o]	91	99	99	109
[u]	83	83	83	86

Fonte: Elaboração própria com base em dados de Balduino (2022) para o PST e o PP e de Escudero et al. (2009) para o PE.

Comparando os números referentes à duração nas variedades de português (PE, PST, PP) e em KVP, de maneira geral o KVP é mais próximo do PP e do PE, o que pode ser explicado pelo fato de o PP ser uma das línguas faladas na ilha do Príncipe, com os falantes de kabuverdiano em geral sendo em sua maioria bilíngues (kabuverdiano-português). Já a maior proximidade com o português europeu poderia ser decorrente do fato de essa ser a norma de prestígio no arquipélago, sendo inclusive a referência alvo da escolarização, inclusive nos materiais didáticos. Essa maior proximidade do KVP com relação ao PP e PE deve ainda ser relativizada em virtude de serem línguas diferentes com sistemas fonológicos também diferentes, sendo possível apontar tendências, mas não uma separação estrita entre o KVP e o PST, por exemplo. Assim, a discussão sobre a duração precisa ser ampliada com um maior número de dados para que traga conclusões mais fidedignas.

## 5 CONSIDERAÇÕES FINAIS

Este artigo apresentou uma descrição acústica das vogais fonéticas orais do kabuverdianu do Príncipe. É possível defender que existe uma variedade principense de kabuverdianu (KVP), que se diferencia daquela(s) encontrada(s) no arquipélago de Cabo Verde, possivelmente como resultado de uma ecologia linguística diversa. Em Cabo Verde, o kabuverdianu é a língua materna da maior parte da população e a mais usada nas interações cotidianas, ao passo que o português, apesar de ser língua oficial, é falado por cerca de 30% da população segundo dados do site Ethnologue. Já no Príncipe, o português é a língua mais falada pela população (conforme dados do Censo de 2012), enquanto o kabuverdianu está mais difundido na comunidade caboverdiana (e seus descendentes), sendo possível conjecturar que essa variedade receberia mais influências do português (dada sua maior difusão), aspecto que precisa ser melhor analisado em estudos futuros.

A partir dos valores do primeiro e do segundo formantes em Hertz, percebeu-se que as diferenças entre as vogais do KVP foram menores do que as observadas em variedades do português, como o PP, o PST e o PE. Essa diminuição nas diferenças pode se refletir em um menor espaço acústico e, por conseguinte, culminar em variação – uma hipótese que precisa ser averiguada a partir da percepção dos falantes de KVP.

No que tange às vogais médias e centrais, ainda que não sido possível determinar o estatuto fonológico dos segmentos, foi possível deslindar que os diferentes segmentos coocorrem nos mesmos contextos, não tendo sido encontrado um condicionamento segmental para um ou outro segmento. Em suma, esta pesquisa, que sugere a relevância da dispersão acústica para a definição do quadro vocálico do KVP, permitiu conhecer um pouco das vogais do kabuverdianu do Príncipe em seus aspectos fonéticos e ampliar o conhecimento sobre a ecologia linguística de São Tomé e Príncipe. As questões levantadas, que possuem pontos de contato com a fonologia, não puderam ser analisadas com maior profundidade nesta pesquisa dada a limitação dos dados e foram apenas sugeridas como aspectos a serem abordados em desdobramentos futuros. Um desses desdobramentos se relaciona às realizações gradientes das vogais, mencionadas de forma breve em nossa análise, tema que demonstra o contato entre a fonética e a fonologia e que certamente trará contribuições relevantes para a pesquisa. As questões referentes à duração também devem ser ampliadas em estudos futuros, contribuindo para a descrição dessa variedade.

## REFERÊNCIAS

- Agostinho AL. Fonologia e método pedagógico do Lung'Ye [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2015.
- Agostinho AL. Sound variation in Portuguese-speaking Africa. In: Zampaulo A, editor. *The Routledge Handbook of Portuguese Phonology*. No prelo.
- Balduino AM. Fonologia do português de São Tomé e Príncipe [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2022.
- Balduino AM, Freitas S. Stressed vowels in São Tomé and Príncipe Portuguese (STPP): acoustic space (F1 and F2) and [ə] production. No prelo.
- Bandeira M, Freitas S. Trabalho de campo em São Tomé e Príncipe. 2018. Inédito.
- Bandeira M. Reconstrução fonológica e lexical do protocioulo do Golfo da Guiné [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2017.

- Baptista MCS. Descrição fonológica da variedade da ilha de Santo Antão. In: Lang J, editor. A variação geográfica do crioulo caboverdiano. Erlangen: FAU University Press; 2014. p. 183-250.
- Barbosa PA, Madureira S. Manual de fonética acústica experimental: aplicações a dados do português. São Paulo: Cortez; 2015.
- Boersma P, Weenick D. Praat: doing phonetics by computer [computer program]. Version 6.2.06. Amsterdam: Universiteit van Amsterdam; restaurada em 23 de janeiro de 2022. Disponível em: <https://www.fon.hum.uva.nl/praat/>.
- Cabo Verde. Decreto-lei nº 67/98. 31 de dezembro de 1998. Aprova, a título referências bibliográficas 355 experimental, o Alfabeto Unificado para a Escrita da Língua Caboverdiana. 5.º Suplemento. I série, n.º 48.
- Carreira A. Migrações nas ilhas de Cabo Verde. Praia: Instituto Cabo-verdiano do livro; 1983.
- Cristófaros Silva T, et al. Fonética Acústica: os sons do português brasileiro. São Paulo: Editora Contexto; 2019.
- Demolin D. The experimental phonology. Revista da ABRALIN. 2012;11(1).
- Escudero P, et al. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. The Journal of the Acoustical Society of America. 2009;126(3):1379-1393.
- Eyzaguirre PB. Small farmers and estates in São Tomé, West Africa [tese]. Yale: Faculty of the Graduate School, Yale University; 1986.
- Feio JA. Cabo-Verdianos e São-tomenses de ascendência cabo-verdiana em São Tomé e Príncipe na atualidade: uma abordagem etnográfica. In: Évora I, organizadora. Diáspora cabo-verdiana: temas em debate. Lisboa: CEsa (Centro de Estudos sobre África, Ásia e América Latina) / ISEG (Instituto Superior de Economia e Gestão); 2016. p. 200-226.
- Flemming E. Contrast and perceptual distinctness. In: Hayes B, Kirchner R, Steriade D, editores. Phonetically based phonology. Cambridge: Cambridge University Press; 2004. p. 232-276.
- Freitas S, Bandeira M, Agostinho AL. Aspectos fonético-fonológicos do kabuverdiano falado na Ilha do Príncipe. Journal of Ibero-Romance Creoles. 2021;11:109-135.
- Freitas S. Vogais pretônicas do kabuverdiano do Príncipe. Em preparação.
- Frota S, Vigário M. Aspectos de prosódia comparada: ritmo e entoação no PE e no PB. Actas do XV Encontro Nacional da APL; 2000. p. 533-555.
- Gleitman, LR, Newport EL. The Invention of Language by children: Environmental and biological influences on the acquisition of Language. In: Gleitman LR, Liberman M, editores. Language: An invitation to cognitive science. 2nd ed. The MIT Press; 1995. p. 1-24.
- Goldsmith JA, editor. The Handbook of Phonological Theory. Wiley; 1996.
- Graham S, Graham T. West Africa lusolexed Creoles word list file documentation; 2014. [citado 02 jul. 2015]. Disponível em: <http://www.sil.org/silest/2004/silest2004-012.html>.
- Hagemeijer T. Initial vowel agglutination in the Gulf of Guinea creoles. In: Enoch A, Smith N, editor. Complex processes in new languages. Amsterdam: John Benjamins Publishing Company; 2009. p. 29-50.
- Hall KC. A typology of intermediate phonological relationships. The Linguistic Review. 2013;30(2):215-275.
- Hualde JI. Quasi-phonemic contrasts in Spanish. In: Schmeiser B, et al., editores. WCCFL 23 Proceedings. Somerville, MA: Cascadilla Press; 2014.
- Instituto Nacional de Estatística (INE). População segundo línguas faladas, RGPH; 2012. [citado 110 set. 2021]. Disponível em: <https://www.ine.st/phocadownload/userupload/Documentos/Atlas/Caracter%C3%ADsticas%20Educativas/5.%20Popula%C3%A7%C3%A3o%20seg.%20L%C3%ADnguas%20Faladas%20-%2054.pdf>.

- Kingston J. The phonetics-phonology interface. In: De Lacy P, editor. *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press; 2007. p. 401-434.
- Ladefoged P. *A course in phonetics*. 2. ed. New York: Harcourt Brace Jovanovich Publishers; 1982 [1975].
- Lang J, editor. *Dicionário do crioulo de Santiago (Cabo Verde) com equivalências de tradução em alemão e português, elaborado por Martina Brüser e André dos Reis Santos (Cabo Verde), com a contribuição de Ekkehard Dengler e Andreas Blum, sob a direcção de Jürgen Lang*. Tübingen: Narr; 2001.
- Lang J. Gramática do crioulo da ilha de Santiago (Cabo Verde). Cap. 1; 2002. [citado 12 nov. 2013]. Disponível em: <http://iilp.wordpress.com/2012/11/14/umagramatica-do-crioulo/>.
- Marusso A. Variabilidade e dispersão vocálica em Português Brasileiro e Inglês Britânico: um estudo de caso. *ALFA: Revista de Linguística*. 2016;60(1):175-201.
- Mongelo JV, Seara IC, Agostinho AL. *Estudo acústico sobre as vogais do guarani: a variedade nhandewa*; 2021.
- Moreira AKT. Descrição da variedade da ilha do Maio. In: Lang J, editor. *A variação geográfica do crioulo caboverdiano*. Erlangen: FAU University Press; 2014. p. 99-180.
- Moreira AKT. *Documentação e descrição gramatical e lexical do crioulo afro-português da ilha do Fogo (República de Cabo Verde, África Ocidental)*. [tese]. Paris: Institut National des Langues et Civilisations Orientales (INALCO), Langage, langues et cultures d'Afrique (LLACAN), National Centre for Scientific Research (CNRS); 2020.
- Nascimento A. Escravatura, trabalho forçado e contrato em S. Tomé e Príncipe nos séculos XIX-XX: sujeição e ética laboral. *Africana Studia*. 2004;7:183-217.
- Nascimento A. *O fim do caminho longi*. Cabo Verde: Ilhéu Editora; 2007.
- Nascimento A. *Vidas de S. Tomé segundo vozes de Soncente*. Cabo Verde: Ilhéu Editora; 2008.
- Nascimento A. *A História da Ilha do Príncipe*. Oeiras: Município de Oeiras; 2010.
- Quint N. *Grammaire de la langue Cap-Verdienne*. Paris: L'Harmattan; 2000.
- Quint N. Vowels as a morphological tool in Santiago Creole Portuguese (Cape Verde). *Journal of African Languages and Linguistics*. 2001;22:69-80.
- Rodrigues U. *Fonologia do caboverdiano: das variedades insulares à unidade nacional* [tese]. Brasília: Instituto de Letras, Universidade de Brasília, Brasília; 2007.
- Santiago AM, Balduino AM. A língua portuguesa em São Tomé e Príncipe: pluricentrismo, colonialidade e ensino. *Domínios de Língu@gem*. 2023;17:e1759.
- Santiago AM, et al. As Vogais no Português do Príncipe. In: Hagemeyer T; Oliveira MSD; Figueiredo C, organizadores. *O Português na África Atlântica*. v. 2; 2022.
- Seibert G. Colonialismo em São Tomé e Príncipe: hierarquização, classificação e segregação da vida social. *Anuário Antropológico/2014*. 2015;40(2):99-120.
- Semedo CIC. *Ilusões do Contrato? Migrações sul-sul, evocações do tráfico, contranarrativas e socialidades dos cabo-verdianos nas roças de São Tomé e Príncipe*. [tese]. Rio de Janeiro: Museu Nacional/Universidade Federal do Rio de Janeiro, Rio de Janeiro; 2016.
- Velupillai V. *Pidgins, creoles & mixed languages: an introduction* [Creole Language Library, 48]. Amsterdam / Philadelphia: John Benjamins; 2015.
- Wendel A, Kaplan A, Jackson S. High functional load inhibits phonological contrast loss: a corpus study. *Cognition*. 2013;128:179-186.
- Wetzels L. Mid Vowel Neutralization in Brazilian Portuguese. *Cadernos de Estudos Linguísticos*. 1992;23:18-55.

## APÊNDICE - Dados de tônicas orais

[i]	[e]	[ɛ]	[ə]	[ɐ]	[a]	[ɔ]	[o]	[u]
bariga	pretu	pé	kiziká	tras	animal	pamodi	katchoru	sukur
minis	skesi	pega	respirar	tempural	tras	korta	vindoris	sukuru
fitcha	medu	bebi	matchu	fala	pasa	mori	vindor	tchuba
kaí	verdi	seku	fla	kaza	kaza	obu	vindô	azul
fika	kabelu	tera	nasi	kazaku	baru	rostu	korta	nuven
dia	kabesa	keba	dimas	mai	brasu	fodja	gosta	suju
obidu	jelu	nebua	brigá	pai	mata	rotcha	obidu	gordura
(a)li	bebi	mudjer	fa	faka	galhu	boka	odju	tuđu
maridu	puera	panela	koraji	fa	pásaru	nobu	poku	kusa
pariba	seka	medu	maw	papá	fumar	pó	oji	susu
tira	mantega	tetu	mata	dimás	kansadu	pikinoti	obi	fruta
riba	sebu	perna	sal	sapu	sidadi	kobra	mordi	pura
briga	pexi	kabesa	pai	kuatu	faka	sol	osu	kurtu
naris	berdi	kanela	dá	mata	pai	nhos	folgu	xuju
ri	djuelu	pertu	lá	animal	koraji	nós	flogô	durmi
vivi	kutubelu	bedju	shupar	sidadi	kasa	forma	obu	juga
figado	rega	kré	sta	karni	agua	odja	forti	lua
kumida	diréta	brumedju	matu	dá	fradu	bó	porku	kuspi
bitchu	iskerda	retu		lá	sapu	dor	fogu	tchupa
raís	juelu	pele		fla	kuatu	omi	otu	lumi
midju	kêl'	pedra		laba	umidadi	nomi	flor	fuma
dividi	stretu	amarelu		estrada	pisadu	fomi	rostu	
mira	rubera	sertu		aza	sabi		piodju	
(a)mi	sfrega	sfrega		pistana	panha		piskos	
vira	area				tomati		noti	
kompridu	sera				karni		pesoa	
vumita	tchera				dá		tchobi	
mininu	strela				babá		kola	
kapina	madera				mai		dos	
nubrina	pretu				estrada		otra	
lubrina	skerda				sal		otru	
fnu	skrebi				fla		kosa	
	estretu				batata		dor	
	eli				fiadu		spoza	
	kapuera				jeladu		gosta	
	kema				masa			
	femia				tchupar			
	tene				nada			
	serenu				rabu			
	teni				lá			

# O onset no angolar moderno

## *The onset in modern Angolar*

Manuele Bandeira\*

*Universidade da Integração Internacional da Lusofonia Afro-Brasileira*, São Francisco do Conde, BA, Brasil

Ubiratã Kickhöfel Alves\*\*

*Universidade Federal do Rio Grande do Sul*, Porto Alegre, RS, Brasil

**Resumo:** O angolar é uma das línguas autóctones da República de São Tomé e Príncipe. Este estudo tem como finalidade descrever o ataque silábico do angolar a partir da Teoria da Sílabas (Selkirk, 1982). Como *corpus*, analisamos 1524 itens coletados, em 2014 e em 2018, em São Tomé e Príncipe, a partir de gravações com bilíngues (falantes de angolar como língua materna e de português como segunda língua). A partir do *software* Dekereke (Casali, 2022), registramos as seguintes estruturas: V, CV, CGV, CCV, CVC, VC e VG. No ataque silábico, é possível haver uma ou duas consoantes, o que pode gerar um *onset* simples de C ou complexo C<sub>1</sub>C<sub>2</sub>, com as seguintes configurações: 1) C<sub>1</sub> pode ser uma obstruente (b, p, t, d, k, g, f, v), seguida por C<sub>2</sub> (l, r); 2) C<sub>1</sub> pode ser uma consoante (b, p, t, d, k, g, f, v, s), seguida por G (w, j). Nasais em início de palavra seguidas por obstruente foram consideradas oclusivas pré-nasalizadas, constituídas por um único segmento de contorno em *onset*. Ademais, constatamos que o ataque silábico no angolar atende ao Princípio da Sequência de Sonoridade, sendo, portanto, crescente até o núcleo silábico: C<sub>1</sub>C<sub>2</sub> (0 > 2); e C<sub>1</sub>G (0 > 3).

**Palavras-chave:** Onset. Angolar. Teoria da sílaba. São Tomé e Príncipe.

**Abstract:** Angolar is one of the autochthonous languages of the Republic of São Tomé and Príncipe. This study aims to describe the syllable onset in Angolar, based on Syllable Theory (Selkirk, 1982). As for the corpus, we analyzed 1524 items collected in 2014 and in 2018 in São Tomé and Príncipe, based on recordings with bilinguals (speakers of Angolar as their mother tongue and Portuguese as their second language). Using the Dekereke software (Casali, 2022), we detected the following structures: V, CV, CGV, CCV, CVC, VC and VG. In syllable onset, one or two consonants are legal. This can generate a simple onset (C) or a complex C<sub>1</sub>C<sub>2</sub> sequence with the following configurations: 1) C<sub>1</sub> may be an obstruent (b, p, t, d, k, g, f, v) followed by C<sub>2</sub> (l, r); 2) C<sub>1</sub> may be a consonant (b, p, t, d, k, g, f, v, s) followed by G (w, j). Word-initial nasals followed by obstruents were considered pre-nasalized stops, consisting of a single contour segment in onset position. Furthermore, we highlight that syllable onsets in Angolar follow the Sonority Sequence Principle, as its elements increase in sonority up to the syllable nucleus: C<sub>1</sub>C<sub>2</sub> (0 > 2); and C<sub>1</sub>G (0 > 3).

**Keywords:** Syllable onset. Angolar. Syllable Theory. São Tomé and Príncipe.

## 1 INTRODUÇÃO

O angolar é uma língua nativa da República de São Tomé e Príncipe (cf. Figura 1) e, ao lado do santome e do lung'le, tem, como ancestral, o protocrioulo do Golfo

---

\* Professora Adjunta, Programa de Mestrado em Estudos de Linguagens, Instituto de Humanidade e Letras, Universidade da Integração Internacional da Lusofonia Afro-Brasileira, São Francisco do Conde, BA, Brasil; [manuelebandeira@unilab.edu.br](mailto:manuelebandeira@unilab.edu.br)

\*\* Professor Associado, Programa de Pós-Graduação em Letras, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil; [ukalves@gmail.com](mailto:ukalves@gmail.com)

da Guiné (doravante PGG) (Bandeira, 2017). O presente artigo tem como objetivo tratar do ataque silábico do angolar (código ISO 639-3: o angolar (AOA)) a partir da Teoria da Sílabas (Selkirk, 1982).



Fonte: Mapmaker (2024).

Figura 1 - Mapa de São Tomé e Príncipe.

Atualmente, o angolar é utilizado como língua nas comunidades costeiras da ilha de São Tomé. Estima-se que um total de 11.377 pessoas falam o angolar como língua materna ou língua adicional, de acordo com dados do INE (2013), em uma população absoluta de 173.015 habitantes do país. Os falantes de angolar estão concentrados em zonas litorâneas específicas, incluindo o distrito de Caué, da Ribeira Afonso até Porto Alegre, e o litoral noroeste, a partir de Neves até Bindá, no distrito de Lembá. Além disso, existem pequenos grupos de falantes próximos à cidade de São Tomé, em São João da Vargem, Pantufo e Praia Melão (Ceita, 1991).



Fonte: Maurer (2013).

Figura 2 - Localização de três comunidades de falantes do angolar em São Tomé.

Para abordar a origem do angolar, é essencial primeiro entender o contexto da colonização portuguesa em São Tomé. No início do século XVI, a interação entre colonos portugueses e escravizados africanos levou ao surgimento do protocioulo do Golfo da Guiné (PGG) (Bandeira, 2017). Em meados do século XVI, alguns falantes do PGG fugiram dos engenhos de açúcar e formaram comunidades isoladas, nas quais

o protocioulo se ramificou no que hoje é conhecido como o angolar, devido a um certo nível de isolamento e à chegada de novos escravizados.

Ferraz (1974) sugere que o angolar está relacionado ao santome e ao lung'le, descartando a hipótese de uma população pré-portuguesa em São Tomé. Ele propõe que os escravizados fugitivos estavam familiarizados com o PGG, o que facilitou o desenvolvimento do angolar, que compartilha 67% do léxico com o santome. A mudança na área de recrutamento de escravizados do Benim para regiões bantu, como o Congo e Angola, trouxe línguas como o kikongo e o kimbundu para a ilha de São Tomé. No entanto, Hagemeyer (2009) argumenta que o kimbundu atuou como uma língua de adstrato, sem efeito fundador, sendo o edo a principal língua de substrato na formação do angolar.

Ainda que o angolar derive do PGG e tenha o substrato edoíde, a influência bantu é evidente em seu léxico. Prova disso é a existência de palavras bantu no angolar, não encontradas no santome. Maurer (1992) afirma que 65% do léxico do angolar é português, 1% de línguas kwa, e 14% de línguas bantu, com 20% de origem incerta, provavelmente bantu. Hagemeyer (2009) confirma que 92% do léxico africano no angolar é bantu, com pequenas contribuições de línguas edoídes e iorubá. Em resumo, o angolar, língua-filha do PGG, já se encontrava constituído quando houve a chegada massiva de escravizados das regiões bantu, cuja influência linguística foi significativa, mas posterior à formação inicial da língua (Bandeira, 2017).

Uma parcela de 6% da população total de São Tomé e Príncipe (STP) utiliza o angolar (INE, 2013). Seus falantes definitivamente não costumam se enquadrar no perfil do monolinguismo, uma vez que possuem um domínio tanto do angolar quanto do português. Essa característica bilíngue implica, obviamente, a influência mútua entre as duas línguas no cotidiano. Entre os utentes de angolar, os mais jovens tendem a falar português como primeira língua (L1) e o angolar como segunda língua (L2), enquanto os mais velhos geralmente têm o português como segunda língua (L2) e o angolar como L1. Como esperado, o contínuo contato com o português leva a variações fonéticas, fonológicas e lexicais no angolar, devido à influência do português, a língua majoritária. Por conseguinte, distinguir entre falantes de angolar como L1 e L2 é crucial para entender as mudanças da língua. Os falantes de português como L1 podem revelar uma maior influência do português no angolar. Em vista disso, nesta pesquisa, analisamos os dados de falantes do angolar como língua materna, procurando assim compreender as suas características em seu uso ordinário (ver seção 4).

Apesar de sua relativa vitalidade, o angolar não tem sido alvo de estudos sistemáticos, tendo a descrição pioneira sido publicada há mais de duas décadas, intitulada *L'angolar: Un créole afro-portugais parle à São Tomé. Notes de grammaire, textes, vocabulaire* (Maurer, 1995). Posteriormente, uma pesquisa de doutoramento foi realizada: *The Angolar creole Portuguese of São Tomé: its grammar and sociolinguistic history* (Lorenzino, 1998). Quase três décadas depois, para realizar a reconstrução do protocioulo do Golfo da Guiné, a língua ancestral dos crioulos nativos de STP, Bandeira (2017) realizou uma descrição do sistema fonológico do angolar. Tal análise fomentou um estudo subsequente com o intuito de estabelecer comparações com as duas propostas pioneiras e propor um estabelecimento preliminar de aspectos fonético-fonológicos gerais a respeito dos sistemas consonantal, vocálico e silábico do angolar moderno (Bandeira; Agostinho; Freitas, 2021).

A pesquisa aqui empreendida busca oferecer uma contribuição aos estudos dos crioulos do Golfo da Guiné. Trata-se de uma análise inédita, tendo em vista que ainda não houve uma pesquisa dedicada somente à descrição do ataque do angolar. É fundamental investigar as características silábicas de uma língua cuja descrição apresenta lacunas. A investigação proposta é, portanto, original e se concentra em aspectos silábicos pouco explorados do angolar, fornecendo uma base para comparação futura com outras línguas africanas, como as línguas dos grupos edo e bantu.

Tendo delineado um breve panorama socio-histórico do angolar, com o intuito de nortear esta pesquisa, foram estabelecidos os seguintes objetivos: i) descrever os padrões silábicos presentes em angolar quanto ao *onset*; ii) identificar as relações entre o ataque silábico do angolar e os Princípios da Sonoridade; iii) comparar a análise da amostra coletada com as descrições prévias de Maurer (1995), Lorenzino (1998), Bandeira (2017) e Bandeira et al. (2021). Na próxima seção, serão observados os padrões silábicos estabelecidos por estudos anteriores sobre a língua.

## 2 ANGOLAR: ESTUDOS PRÉVIOS SOBRE A SÍLABA

O inventário fonológico consonantal do angolar é constituído por dezesseis fonemas: /p, b, t, d, k, g, f, v, θ, ð, l, m, n, ɲ, j, w/ (Bandeira, 2017; Bandeira, Agostinho, Freitas, 2021). Quanto ao seu sistema vocálico, o angolar possui sete vogais orais (/i e ε a o u/) e sete longas (/ii ee εε aa ɔɔ oo uu/), de acordo com Bandeira (2017). A respeito da sílaba do angolar, trataremos, primeiramente, das propostas pioneiras, a primeira feita por Philippe Maurer (1995) e a segunda por Gerardo Lorenzino (1998). No Quadro 1, estão demonstradas as estruturas silábicas encontradas em quatro estudos:

Quadro 1 – Estruturas silábicas do angolar segundo Maurer (1995), Lorenzino (1998), Bandeira (2017) e Bandeira et al. (2021).

	Maurer (1995)	Lorenzino (1998)	Bandeira (2017)	Bandeira, Agostinho e Freitas (2021)
V	<b>a</b> <i>PRO</i>	<b>a</b> <i>PRO</i>	<b>o</b> 2PL	<b>o</b> 2PL
VV	<b>ôu</b> ‘costurar’	---	<b>ee</b> ‘sim’	<b>ee</b> ‘sim’
CV	<b>pê</b> ‘pôr’	<b>ta</b> <i>TMA</i>	<b>bithu</b> ‘bicho’	<b>bithu</b> ‘bicho’
CVV	<b>fira</b> ‘ferida’	<b>beega</b> ‘barriga’	<b>faaku</b> ‘fraco’	<b>faaku</b> ‘fraco’
CGV	---	<b>kwa</b> ‘coisa’	<b>thagwa</b> ‘enxaguar’	<b>thagwa</b> ‘enxaguar’
VN	<b>anda</b> ‘andar’	---	<b>andhu</b> ‘recém-nascido’	<b>andhu</b> ‘recém-nascido’
CVN	---	---	<b>lamba</b> ‘cubrir’	<b>lamba</b> ‘cubrir’
CVVN	---	---	<b>laandha</b> ‘laranja’	<b>laandha</b> ‘laranja’
CGVN	---	---	<b>kwanda</b> ‘cume’	<b>kwanda</b> ‘cume’
N	---	---	---	<b>Nge</b> ‘ser humano’
CCV	---	---	---	<b>bixikleta</b> ‘bicileta’

Fonte: Bandeira (2017); Bandeira, Agostinho, Freitas (2021).

De acordo com a análise dos dados coletados e com as descrições de Maurer (1995) e Lorenzino (1998), o angolar apresenta padrões fonológicos silábicos como V, VV, CV, CVV, CGV e VN.

Bandeira (2017), em sua tese de doutoramento, tinha como objetivo primordial a reconstrução fonológica e lexical do protocrioulo do Golfo da Guiné a partir de comparações sistemáticas de suas línguas-filhas, o angolar, o santome, o lung<sup>2</sup>Te e o fa

d'Ambô. Diante disso, o *corpus* do angolar utilizado para análise esteve circunscrito aos cerca de 540 conjuntos de cognatos, pertencentes ao léxico básico e mais antigo das quatro línguas. Na análise, Bandeira (2017) observou que o angolar apresentava os seguintes tipos silábicos: V, VV, CV, CVV, CGV, VN, CVN, CVVN e CGVN. Segundo a autora, o angolar sistematicamente apresentava dissoluções de *onsets* complexos de itens herdados da sua língua-mãe, como se pode ver no Quadro 2:

Quadro 2 - Dissolução dos *onsets* complexos do protocioulo no angolar antigo.

Protocioulo	Angolar	Glosa
* <b>blasu</b>	[ˈba:θu]	‘braço’
* <b>blaNku</b>	[ˈbɐ:kɔ]	‘branco’
* <b>gleza</b>	[ˈge:ðɐ]	‘igreja’
* <b>pletu</b>	[ˈpe:tɔ]	‘preto’
* <b>plaga</b>	[ˈpa:ɣɐ]	‘praga’
* <b>sɔ'pla</b>	[sɔ'pa]	‘assoprar’

Fonte: Bandeira (2017).

Contudo, em sua tese, Bandeira (2017) já alertava para a existência de uma flutuação nos padrões silábicos, em um recorte sincrônico, a partir do apagamento da líquida com ou sem alongamento compensatório, por um lado, e a presença de *onsets* complexos, por outro:

1. [ˈgo:su] ~ [ˈglosu] ‘grosso’
2. [ˈfo:ri] ~ [ˈflori] ‘flor’
3. [obli'ga] ‘obrigar’
4. [bu'ga] ~ [blu'ga] ‘descascar’
5. [ma'de] ~ [ma'dre] ‘útero’

Diante do quadro apresentado por Bandeira (2017), permaneceu como ponto em aberto a seguinte pergunta: 1- existe *onset* complexo no angolar moderno? Em vista desse questionamento sobre sílaba e sobre outros pontos, Bandeira, Agostinho e Freitas (2021) apresentam um estudo sobre aspectos fonético-fonológico do angolar moderno com base nos dados reunidos na tese de doutoramento de Bandeira (2017). Para a pergunta supramencionada, as autoras apresentam resposta afirmativa. De acordo com as autoras, o *onset* pode consistir em um ou dois segmentos, incluindo C, CC e CG, em que a segunda consoante pode ser /j/, /w/, /l/ ou [r]<sup>2</sup>. A resposta afirmativa para a pergunta supramencionada levanta questões que merecem futuras investigações. Possíveis influências externas, como o contato com o português, poderiam ter contribuído para o desenvolvimento atual do angolar, como a presença de *onsets* complexos. Contudo, ainda que as variedades urbanas do português de São Tomé e Príncipe demonstrem um alto índice de dissolução de *onsets* com essa configuração (Balduino, 2022), não se sabe ao certo a respeito do comportamento dos ataques complexos em variedades faladas em zonas interioranas. Não se pode, por

<sup>2</sup> Bandeira, Agostinho e Freitas (2021) identificaram variações lexicais entre [r], [d] e, ocasionalmente, [l], mas não encontraram pistas suficientes para classificar o rótico como um fonema independente no angolar. Segundo as autoras, pares mínimos mostraram /d/ e /l/ como fonemas distintos (por exemplo, **data** ‘quantidade’ versus **lata** ‘lata’), enquanto [r] sempre variou com [d] e às vezes com [l]. Em alguns casos, como **dooba** ‘dobra’ (**rooba** não é aceito), [d] e [r] não são intercambiáveis sem que o item seja reconhecido como fora do angolar. Informantes mais velhos associam algumas variações com [d] ao português, indicando uma influência externa.

outro lado, descartar a possibilidade de que resquícios históricos do PGG tenham sido preservados na estrutura fonológica do angolar moderno. As autoras assumem a mesma posição de Agostinho (2015, 2016) para o lung'Ie e de Araujo e Agostinho (2014) para o fa d'Ambô em relação aos glides, porque as três línguas são geneticamente relacionadas e os glides se comportam como consoantes no jogo de linguagem em lung'Ie e fa d'Ambô<sup>3</sup>. Haveria, portanto, as seguintes estruturas silábicas em angolar, de acordo com Bandeira et al. (2021): V, N, VN, CV, GV, CVN, CGV, CGVN, VV, CCV, CVV e CVVN.

De acordo com Bandeira et al. (2021), o núcleo silábico deve ser preenchido por pelo menos uma vogal simples ou por uma nasal silábica, sendo a sílaba mínima V ou N. Segundo Bandeira et al. (2021), como nas outras línguas-irmãs, em angolar, a nasal silábica, na posição de núcleo, seria uma consoante hospedeira que ocorre em conjunto com outra consoante, como em /Nge/ ['ŋ.ge] 'ser humano'. O angolar seria, por conseguinte, a língua-filha do PGG mais resistente ao preenchimento da coda, conforme observado por Maurer (1995), permitindo que apenas a consoante /N/ ocupasse essa posição<sup>4</sup>. No entanto, a análise proposta por Bandeira et al. (2021) suscita uma questão: onde, de fato, está localizada a nasal no angolar: no núcleo ou na coda? Assumir simultaneamente as duas posições para a nasal, como sugerido por Bandeira et al. (2021), não parece um pressuposto consistente, o que demanda uma nova abordagem no que diz respeito à consoante. Ao reconhecer esse fato, torna-se crucial refletir sobre a natureza da suposta nasal silábica no angolar e sua posição na sílaba. Portanto, a presente análise poderá viabilizar um avanço na compreensão acerca da sílaba do angolar. Além disso, esta análise é pioneira ao aplicar o referencial teórico referente à Teoria da Sílaba, a partir da estrutura silábica de Selkirk (1982), na investigação da estrutura silábica do angolar. Na próxima seção, apresentaremos brevemente o nosso referencial teórico sobre a Teoria da Sílaba (Selkirk, 1982).

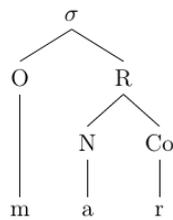
### 3 REFERENCIAL TEÓRICO: TEORIA DA SÍLABA (SELKIRK, 1982)

Os fundamentos que regem a construção da sílaba básica referem-se à hierarquia de importância existente entre os componentes que a constituem, conhecidos como Princípios de Composição da Sílaba Básica, ou seja, entre o ataque e a rima, bem como entre o núcleo e a coda, como em (1) (Selkirk, 1982, Bisol, 2013).

<sup>3</sup> Dado que a relação genética não configura um argumento fonológico capaz de sustentar o comportamento consonantal dos glides no angolar moderno, na presente análise, discutiremos o status dos glides, conforme poderá ser visto na seção 5.

<sup>4</sup> Nos termos do projeto de pesquisa maior a partir do qual foi extraída a presente análise, defendemos a posição de que a coda do angolar possibilita a ocorrência de outras consoantes, tais como as fricativas [s, ʃ], a lateral [l], as nasais [n, ŋ, m] e os glides [j] e [w]. Os resultados específicos relacionados à coda do angolar serão detalhados em um artigo futuro.

(1)



Além da subdivisão entre onsets e rimas, a sílaba pode ser organizada a partir da sonoridade de seus segmentos. Por conseguinte, há uma sonoridade crescente no ataque e decrescente na coda (Bisol, 2013). Assumindo a condição de sequência de sonoridade, tanto o ataque quanto a coda organizam-se em relação ao núcleo, que é o elemento mais sonoro da sílaba. Em vista disso, o ataque é composto por elementos de menor sonoridade comparado ao núcleo. Até o pico, portanto, a sonoridade tende a se encontrar (sobretudo em sequências de duas consoantes em ataque) em escala crescente. A coda, por seu turno, é também composta por elementos de sonoridade inferior ao núcleo silábico, o que leva a uma escala decrescente de sonoridade do pico. O núcleo, portanto, é o componente central da sílaba, cercado por elementos cuja sonoridade segue um padrão de acréscimo ou decréscimo.

A sonoridade é comumente representada através de uma escala que diz respeito a um ordenamento de segmentos que parte dos mais altos em sonoridade, como as vogais, até os mais baixos na sonoridade, como as obstruintes. Assim, temos o seguinte ordenamento adaptado de Zec (2006, p. 177):

(2) **V** (vogais) > **G** (Glides) > **L** (líquidas) > **N** (nasais) > **O** (obstruintes)

Blevins (1995, p. 209) salienta a relação que subjaz entre qualquer membro de uma sílaba e o pico silábico, em que somente sons mais altos na hierarquia de sonoridade são permitidos. A sonoridade é um importante condutor no que diz respeito à constituição silábica de uma língua. Portanto, através da escala de sonoridade, a composição de uma sílaba básica é iniciada com identificação do núcleo. Assim, o pico de sonoridade mais alto na sílaba é o núcleo. A partir desse ponto, a rima é projetada, e a sílaba é constituída. Somente depois da projeção da rima, inicia-se a ramificação à esquerda; desse modo, consoantes adjacentes são mapeadas de acordo com o Princípio da Maximização do Ataque (Selkirk, 1982; Clements; Keyser, 1983). A partir desse ponto, a coda, ao se adjungir à rima das consoantes contíguas até então não silabificadas, pode ser formada.

Ainda sobre princípios universais e condições particulares das línguas, há três generalizações importantes quanto à sonoridade e ao núcleo em conjunto com a Generalização de Sequenciamento Silábico (*Sonority Sequencing Generalization*), a saber: (i) todas as línguas têm sílabas que contêm núcleo vocálico não alto; (ii) se uma língua aceita um segmento silábico com valor de sonoridade  $x$ , então todos os segmentos com valores de sonoridades maiores do que  $x$  são também potenciais núcleos silábicos; (iii) dentro de uma língua, núcleos opcionais nunca são mais sonoros do que os núcleos obrigatórios (Blevins, 1995, p. 216).

Quadro 3 - Escala de soância (O: obstruinte; N: nasal; L: líquida; G: glide; V: Vogal).

O	< N	< L	< G	< V	
-	-	-	-	+	silábico
-	-	-	+	+	vocóide
-	-	+	+	+	aproximante
-	+	+	+	+	soante
0	1	2	3	4	grau de soância

Fonte: Clements (1990, p. 294).

Baseado também em uma escala de sonoridade, Clements (1990) postula o *Princípio de Sequenciamento de Soância* (PSS). O PSS pressupõe que os segmentos mais altos da escala de sonoridade ficam no núcleo da sílaba, ao passo que os segmentos mais baixos da escala de sonoridade deslocam-se para as margens. Clements (1990, p. 294) propõe uma escala, descrita no Quadro 2, que se assemelha àquela proposta por Selkirk (1984), sobrepujando-se àquela, ao acrescentar o grau de soância. De acordo com Ramos e Tenani (2009, p. 25), o acréscimo do grau de soância permitiu que se pudesse explicar a preferência, no português brasileiro, por sílabas com elementos adjacentes em ataques e codas complexos separados por uma distância mínima de dois graus na escala ( $0 < 2 < 4$ ).

De acordo com a Lei do Contato Silábico (Murray; Vennemann, 1983), numa sequência de segmentos heterossilábicos do tipo A\$B, A é preferencialmente mais sonoro do que B (Collishonn, 1999, p. 102). Desse modo, conjugando o PSS e a Lei do Contato Silábico, constata-se que a alta sonoridade é esperada em segmentos em posição de coda, ao passo que a baixa sonoridade é esperada em segmentos em posição de onset. Por conseguinte, sendo a silabificação o processo de mapeamento de uma sequência de segmentos ao molde silábico de uma língua, é necessário que o segmento que se candidata a ocupar uma determinada posição na sílaba siga a ordem de sonoridade crescente em direção ao pico e decrescente a partir dele (Bisol, 2013). Após uma breve apresentação da nossa fundamentação teórica, na próxima seção, trataremos dos procedimentos metodológicos executados para a realização do estudo.

#### 4 METODOLOGIA

Para a execução da pesquisa, foram conduzidas duas viagens de campo na comunidade de São João dos Angolares, em São Tomé, no início de 2014 e no final de 2018. As viagens duraram até dois meses, durante as quais foram feitas gravações com falantes de angolár. Além das entrevistas, que tiveram uma duração média de 30 minutos, foram registradas histórias da comunidade e realizados testes com pares mínimos, julgamentos de gramaticalidade e testes de reação subjetiva para verificar a aceitabilidade de certas realizações e observar aspectos fonológicos da língua.

Durante a coleta, os itens foram submetidos à conferência de grafia, transcrição e significado em conjunto com os falantes. Assim, em caso de erro, o item era prontamente corrigido. Feito isso, foram gravados, com os falantes, itens que não tivessem sido registrados anteriormente, mas que, por meio de listas de palavras em português, fosse possível questioná-los sobre como os mesmos poderiam dizer a mesma ideia ou palavra em angolár. Ao mesmo tempo em que eram feitas as gravações,

foi feita também a transcrição dos dados coletados, com base em uma análise de oitiva. A transcrição, feita imediatamente após as gravações, era de suma importância, uma vez que, em caso de a gravação apresentar uma qualidade de áudio ruim ou mesmo se o dado estivesse inaudível, era possível, sendo qualquer uma dessas situações detectadas, resolver por meio da repetição da gravação do item que ofereceu dificuldades ao entendimento. Os dados coletados foram transcritos e resultaram em uma lista de cerca de 2900 verbetes com transcrições fonéticas, que serviu como base para esta pesquisa.

Embora as viagens de campo tenham sido fundamentais para a consecução do trabalho - viabilizando a formação do *corpus*, trazendo, por conseguinte, uma maior fluidez para análise dos dados -, tivemos alguns contratemplos durante a estadia, devido à infraestrutura precária do país, tais como o constante racionamento ou interrupção do fornecimento de energia elétrica (durando até mesmo cinco dias). Ademais, não encontramos, nas duas pesquisas de campo, um espaço adequado para captação precisa de áudio; por essa razão, fizemos as gravações em ambientes abertos, em geral, na parte externa das casas dos participantes, levando ao descarte de alguns trechos gravados pela impossibilidade de acuidade auditiva.

Para a execução da pesquisa, nos concentramos nos dados coletados em gravações espontâneas (1524 palavras de cerca de 2900 verbetes) para fazer as análises em geral. Diante do intenso contato com o português e pela natureza do objeto de análise do presente estudo, a sílaba, foram considerados os dados da primeira coleta dos falantes que apresentavam o angolar como língua primeira, constituindo, assim, um *corpus* de análise de 1524 itens com 4 participantes bilíngues que têm o angolar como língua materna e o português como L2 (ver quadro 2). Cabe ainda dizer que, para a análise especificamente das nasais + obstruintes em início de palavra, optamos pelos dados reunidos em gravações controladas que tiveram uma duração média de 10 minutos. A escolha pelo *corpus* coletado em gravação controlada com frases-veículo (**N ga fa x bathu / Eu falo x baixo**), ainda que com um número reduzido de itens (140 palavras de 3 participantes), se dá pela natureza do objeto (nasal + obstruinte), posto que é importante isolar a variável e garantir o mesmo item e contexto de realização (como posição na sílaba e na palavra) em casos de aspectos pouco descritos. Os participantes foram instruídos a repetir, três vezes, as frases-veículo da forma mais natural e espontânea possível. Assim, realizamos uma verificação acústica de algumas dessas 140 palavras. Para tanto, utilizamos o software *Praat – version 6.2.20* (Boersma; Weenink, 2022), para observar a consoante nasal antecedente a obstruinte em início de palavras.

No que diz respeito à composição do *corpus*, é válido destacar que o número de participantes englobados na pesquisa é relativamente reduzido (4 quatro), o que, por sua vez, se configura como uma das limitações inerentes a este estudo. Além desse aspecto, é importante mencionar que os participantes selecionados não se enquadram no perfil de monolinguismo, uma vez que possuem um domínio competente tanto do angolar quanto do português. A avaliação desse domínio foi estabelecida por meio da observação em trabalho de campo das interações dos participantes com outros falantes em ambas as línguas.

Ademais, a diferenciação entre os falantes de português como L1 e L2 se mostrou uma variável relevante para a descrição da sílaba. Essa distinção foi estabelecida a partir dos dados coletados por meio de um questionário social aplicado

a cada participante. O questionário abordou, entre outros aspectos, a primeira língua adquirida na infância e a língua utilizada predominantemente na comunicação com o núcleo familiar durante a infância. Não houve relato de aquisição simultânea das duas línguas na primeira infância (ou seja, nos primeiros seis anos completos). Para os 4 participantes que têm o angolar como língua materna, o contato mais intenso com o português se iniciou com a vida escolar (a partir dos 7 anos de idade). O angolar, por estar em contato contínuo com o português, pode apresentar variações na sua manifestação fonético-fonológica e lexical, conforme influências de estrutura da língua majoritária. Atualmente, não é mais necessário haver o deslocamento para os centros urbanos para que os angolares tenham contato com o português. Em comunidades como São João dos Angolares, o português se tornou a língua franca, a língua de ensino e a língua materna da maior parte dos seus residentes. Esse cenário, no entanto, muda em zonas mais interioranas, como a comunidade de Rio Grande, onde as crianças aprendem português na escola aos 6 ou 7 anos, mas, dentro da comunidade, a língua efetivamente utilizada é o angolar. Em contraste, na comunidade de São João dos Angolares, o português já suplantou o angolar em uso cotidiano. Em vista disso, a distinção entre os falantes de angolar como L1 e L2 foi importante para avaliarmos as possíveis implicações na evolução da língua. Aqueles cuja língua materna é o angolar podem proporcionar uma realização mais representativa da estrutura linguística nativa, enquanto os falantes de português como L1 podem oferecer insights sobre a maior influência do português no angolar. Assim, ao concentrarmos nossa análise nos dados da primeira coleta dos falantes do Grupo 1 (indivíduos bilíngues cuja língua materna é o angolar), buscamos capturar as particularidades específicas do angolar em seu vernáculo e minimizar o efeito inevitável do contato com o português.

No Quadro 4, dispomos o quadro social dos participantes do *corpus* efetivamente utilizado.

Quadro 4 – Participantes do *corpus* utilizado (ensino: básico (até 4 anos), fundamental (6-9 anos), secundário (10-12 anos) e graduação).

Falante	Idade	Sexo	L1	L2	Ensino	Ano da coleta	Gravação
2	71	M	Angolar	Português	Básico	2014 (1ª coleta)	Espontânea Controlada
4	32	M	Angolar	Português	Fundamental	2014 (1ª coleta)	Espontânea Controlada
7	33	M	Angolar	Português	Graduação	2014 (1ª coleta)	Espontânea
11	44	M	Angolar	Português	Básico	2018 (1ª coleta)	Espontânea Controlada

Fonte: Elaborado pelos autores.

Como o recorte utilizado foi feito a partir de falantes de angolar como L1, enfrentamos dificuldades consideráveis para reunir participantes com esse perfil, resultando em uma amostra reduzida e desbalanceada. Não obtivemos equilíbrio quanto à idade e ao sexo dos participantes: todos são homens, com uma disparidade etária entre um participante idoso (71 anos) e os demais (intervalo de 32 a 44 anos). Essa discrepância se deve a fatores socioeconômicos e culturais na comunidade de São João dos Angolares. As mulheres enfrentam uma tripla jornada de trabalho, o que limita seu tempo livre e disponibilidade para participar da pesquisa. Além disso, os homens frequentemente consomem álcool em excesso, afetando sua prontidão para

participar de gravações. Para mitigar esses desafios, as gravações foram estrategicamente agendadas após o término do trabalho e antes do início da ingestão alcoólica. Reconhecemos que essas condições podem ter influenciado a amostra e destacamos a necessidade de considerar essas questões em futuras pesquisas para obter uma amostra mais representativa em termos de sexo e idade. Outra discrepância observada diz respeito à educação formal, pois reconhecemos que a formação educacional pode influenciar o contato linguístico dos participantes com o português, especialmente o participante com graduação completa, que provavelmente tem um contato mais intenso com o português. Apesar de o referido desequilíbrio ser uma limitação, consideramos que tal fato não compromete a validade das observações fonético-fonológicas feitas. Para pesquisas futuras, buscaremos uma amostra mais homogênea em termos de escolaridade ou um controle mais rigoroso dessa variável.

Os itens observados foram tabulados e submetidos a um levantamento automático de padrões silábicos a partir do *software* Dekereke (Casali, 2022). O *software* fornece diversas ferramentas para o tratamento de dados, permitindo a investigação de propriedades fonotáticas de qualquer tipo de *corpus*, gerando gráficos de segmentos e sequências fonológicas, como quadros de vogais e consoantes e padrões silábicos. Como os padrões silábicos obtidos, passamos à checagem item a item para averiguar se a atribuição silábica automática atendeu aos princípios teóricos norteadores da análise, como o Princípio de Sequenciamento de Sonoridade (Selkirk, 1982; 1984; Clements, 1990). Após esse exame sistemático dos itens, na próxima seção, os padrões silábicos no angolár serão propostos no que diz respeito ao *onset*.

## 5 RESULTADOS E DISCUSSÃO

De acordo com o resultado obtido pelo levantamento automático de sílabas do Dekereke (Casali, 2022), o angolár apresenta os seguintes tipos de sílabas: V, CV, CGV, CCV, CVC, VC e VG (ver Quadro 5).

Quadro 5 - Resultado geral das sílabas encontradas a partir da realização fonética dos itens.

Tipo Silábico	Início de Palavra	Meio de Palavra	Final de palavra	Exemplos do tipo silábico
CV	1421	634	1546	[d <sub>o</sub> . 'to.l <sub>o</sub> ] “médico”
V	172	56	55	[u. 'fwa] “odor”, [a. 'bli] “abril”
CGV	55	2	56	[ 'kwe.l <sub>o</sub> ] “coelho”, [me. 'dʒja] “meio dia”
CVC/CVG	24	3	13	[a.mis.tra. 'sõ] “administração”, [ 'sef.te 'fe.l <sub>e</sub> ] “sexta-feira”, [mal. 'va.d <sub>o</sub> ] “pessoa ruim”, [ 'mɛ.gãj] “mulher”, [ 'mil] “mil”
CCV	22	0	10	[ 'fla.k <sub>o</sub> ] “fraco”, [fle. 'ge.ʒi] “frequês”, [ 'qla.vi] “bonito/a”
VC	3	0	1	[ 'os.te] “ostra”, [i. 'tu. 'de.ti] “aluno”, [is. 'pa.d <sub>e</sub> ] “espada”, [ 'ʒu.is] “juíz”
VG	1	0	0	[ 'aw.to ka.ru] “ônibus”

Fonte: Elaborado pelos autores.

A partir do Quadro 5, podemos observar que o tipo silábico mais encontrado nos dados foi o CV, seguido por sílaba V. Tal resultado era esperado, tendo em vista a tendência geral das línguas em direção ao tipo CV e também devido à diacronia da

língua. O angolano antigo apresentava, em seu inventário, um favorecimento por sílaba com *onsets* simples e sílabas leves/abertas (isto é, sem rima ramificada).

No ataque silábico, é possível haver uma consoante ou um glide compondo um *onset* simples. Todas as seguintes consoantes podem ocupar a posição de *onset* em início e meio de palavra: /p, t, k, b, d, g, f, θ, s, ʃ, v, ð, z, ʒ, m, n, ɲ, l, r/, com a exceção de /ʎ/, cujos registros não indicaram a realização em posição inicial de palavra.

Quadro 6 - Consoantes constituintes de *onset* simples no angolano.

<i>Onset</i>	<i>Início de Palavra</i>	<i>Meio de Palavra</i>
<i>p</i>	[pe.'ka.rɔ] “pecado”	[su.'pe.tɔ] “sábio”
<i>b</i>	[be.'ku] “tartaruga”	[fe.'be] “febre”
<i>t</i>	[t'a.fɐ] “tarrafa”	[i.'te] “útero”
<i>d</i>	[d'o.lo] “tristeza”	[gũ.dɐ] “seduzir”
<i>k</i>	[kɐ.'dʒã] “candeeiro”	[a.fi.kɐ] “África”
<i>g</i>	[ga.'i.ru] “grande”	[i.'go] “pescoço”
<i>f</i>	[fi.'ni.ni] “funil”	[a.'fa.si] “alface”
<i>v</i>	[vũ.gɔ] “canção”	[kɐ.'vi] “palha do andim”
<i>θ</i>	[θɛ.ða.lɐ] “senzala”	[pe.'θa] “pensar”
<i>ð</i>	[ðõ.'ge] “cesto de lixo”	[pe.'ða] “prestar”
<i>ʃ</i>	[i.'vi] “servir”	[gɔ.'fi] “noite”
<i>ʒ</i>	[ʒi.ʒi.'ka] “gravata”	[fle.'ge.ʒi] “freguês”
<i>m</i>	[ma.'la.ʃɔ] “filhote de tubarão”	[la.mɐ] “lama”
<i>n</i>	[no] “1ª pessoa do plural”	[a.'mõ.ni] “amante”
<i>ɲ</i>	[ɲo.'ko] “cansar”	[i.'ɲɛ] “unha”
<i>l</i>	[la.'zõ] “razão”	[fu.'la] “furar”
<i>r</i>	[ra.'ra] “tipo de peixe”	[de.rɔ] “dedo”
<i>ʎ</i>	---	[ku.'ʎe.ri] “colher”
<i>s</i>	[so.ga] “sogra”	[kɐ.'se.rɐ] “asma”
<i>ʒ</i>	[zu.'ga] “atirar”	[le.'za.rɔ] “manco”

Fonte: Elaborado pelos autores.

O *onset* complexo pode ser constituído por CC, sendo que o segundo segmento pode ser /l/ ou /r/. Quase todas as consoantes obstruintes podem ocupar a primeira posição de *onset* em início e meio de palavra: /p, b, t, d, k, g, f/ diante de /l/ ou [r], com a exceção de /v/, /t/ e /d/, cuja realização não foi observada nos registros com consoante líquida na segunda posição.

Quadro 7 - Consoantes constituintes de *onsets* complexos no angolano, cujo C2 é uma consoante líquida.

C1	C2	Início de Palavra	Meio de Palavra
<b>p</b>	l	[plɛ me.'lõ] “Praia Melão”	[i.'zɛ.plɔ] “exemplo”
<b>b</b>	l	[bla.'bi] “varize”	[tu.'blõ] “tubarão”
<b>t</b>	l	---	---
<b>d</b>	l	---	---
<b>k</b>	l	[kli.'ni.ka] “farmácia”	[bi.'i.kle.tɐ] “bicicleta”
<b>g</b>	l	[gla.vi] “belo”	---
<b>f</b>	l	[fla.kɔ] “frágil”	---
<b>v</b>	l	---	---
<b>t</b>	r	[trĩ.kɔ] “fechadura”	[kõ.trɐ] “amuleto”
<b>g</b>	r	[gra.'va.tɐ] “gravata”	---
<b>d</b>	r	---	[ma.'dri] “útero”

<b>f</b>	ɾ	[ˈfrõ.nɐ] “fronha”	---
<b>p</b>	ɾ	[ˈpɾi.mõ] “primo”	[kõ.pɾi.ˈka] “complicar”

Fonte: Elaborado pelos autores.

Ademais, é possível haver um *onset* complexo constituído também por CG, sendo o segundo segmento /j/ ou /w/.

Quadro 8 - Consoantes constituintes de *onsets* complexos no angolar com glide [w] na segunda posição do *onset*.

C	G	Início de Palavra	Meio de Palavra
<b>b</b>	w	[ˈbwa.rõ] “bom”	[tu.ˈbwa] “tubarão”
<b>d</b>	w	---	[mõ.ˈdwa] “monte pequeno”
<b>f</b>	w	---	[ũ.ˈfwa] “mau cheiro”
<b>v</b>	w	---	[u.ˈvwa] “nove”
<b>s</b>	w	---	[uˈswa] “vinho de palma podre”
<b>g</b>	w	[ˈgwa.rɐ] “guarda”	[sa.ˈgwa] “enxaguar”
<b>k</b>	w	[ˈkwa.ni] “cesto”	[ˈmẽ.kwa] “manco”
<b>ʒ</b>	w	---	[a.me.ˈʒwa] “ameijoia”

Fonte: Elaborado pelos autores.

Quadro 9 - Consoantes constituintes de *onsets* complexos no angolar com glide [j] na segunda posição do *onset*.

C	G	Início de Palavra	Meio de Palavra
<b>p</b>	j	[te.ˈpja] “tempera”	---
<b>t</b>	j	---	[ˈfu.tja] “estrangeiro”
<b>d</b>	j	---	[u.ˈdja] “orelha”
<b>g</b>	j	---	[le.ˈgja] “alegria”
<b>f</b>	j	[ˈfjɐ] “folha”	---
<b>s</b>	j	---	[de.mo.ka.ˈsjɑ] “democracia”
<b>z</b>	j	---	[vi.zja.ˈrõ] “inspetor”
<b>m</b>	j	[ˈmjõ.gɐ] “mar”	[ko.no.ˈmja] “economia”
<b>n</b>	j	---	[a.lu.ˈmi.nju] “alumínio”
<b>r</b>	j	---	[a.lu.ˈma.rju] “armário”

Fonte: Elaborado pelos autores.

No angolar, é possível observar uma flutuação na realização dos ditongos crescentes<sup>5</sup> (uma vogal antecedida por [j] ou [w]). Tais ditongos podem ser realizados de duas formas: como ditongo ou como hiato, como em (1) a (4).

- (1) [ˈbwa.rõ] ~ [ˈbu.a.rõ] “bom”
- (2) [ũ.ˈfwa] ~ [ũ.ˈfu.a] “mau cheiro”
- (3) [ˈmjõ.gɐ] ~ [mi.ˈõ.gɐ] “mar”
- (4) [vi.zja.ˈrõ] ~ [vi.zi.a.ˈrõ] “inspetor”

Como ocorre no português brasileiro, os ditongos crescentes no angolar podem ser considerados hiatos lexicais. Isso significa que os ditongos crescentes são derivados pós-lexicalmente por meio da ressilabificação: **bu.a.ru** se torna **bwaru**

<sup>5</sup> Em função do caráter de delimitação deste estudo, não apresentaremos, aqui, os espectrogramas referentes a cada uma dessas formas. Em um estudo futuro, referente, especificamente, à estrutura dos glides, abordaremos as produções acústicas dessas formas variantes.

“bom” e **mi.on.ga** se torna **mjon.ga** “mar”. Assim, em nível pós-lexical, a vogal alta pode se transformar em glide, formando assim o ditongo crescente<sup>6</sup> (Bisol, 2013, p. 40).

Considerando os dados até aqui apresentados, o ataque silábico no angolar atende ao Princípio da Sequência de Sonoridade (O- obstruente: 0; L - líquida: 2; G-glide: 3), sendo, portanto, crescente até o núcleo silábico: C<sub>0</sub>1C<sub>L</sub>2 (0 < 2); e C<sub>0</sub>G (0 < 3).

### 5.1 Sobre a nasal e a obstruente

A similaridade articulatória entre obstruintes e nasais pode levar a uma conexão na fonologia de várias línguas, em que sequências de nasal seguida por uma oclusiva devem ser homorgânicas. Em certos casos, tais sequências são tratadas como segmentos unitários, especialmente quando ocorrem no início da sílaba, sendo denominadas “oclusivas pré-nasalizadas”. Esse fenômeno é indicado por um símbolo nasal sobrescrito antes do símbolo da oclusiva (Ladefoged; Maddieson, 1996, p. 199). Em angolar, foram encontradas realizações de nasais e obstruintes em início de palavra, em que a consoante nasal assume o ponto de articulação da oclusiva imediatamente posterior, como no Quadro 10.

Quadro 10 - Nasal + Obstruente em início de palavra em angolar.

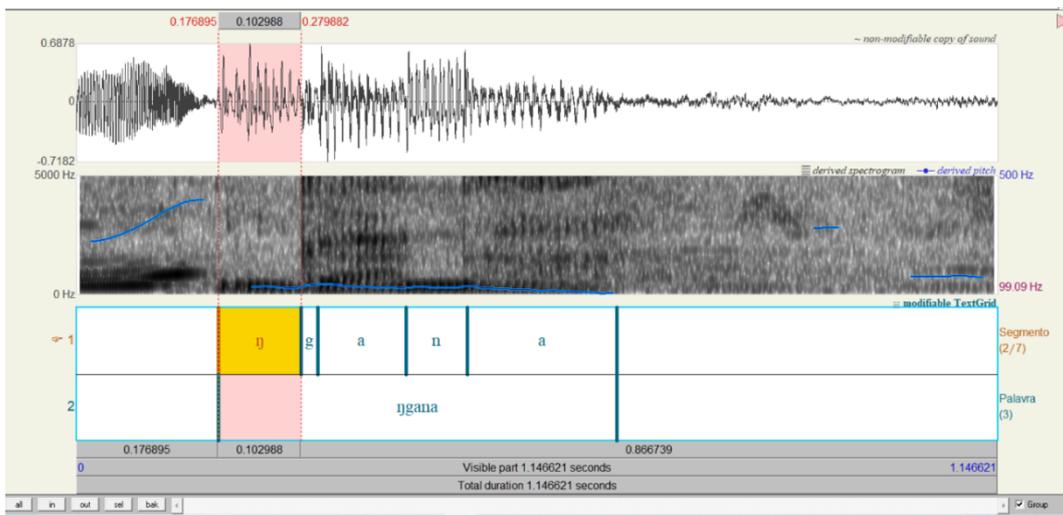
C	C	Início de Palavra
<b>n</b>	t	[nta.'ri] “metade”
<b>n</b>	d	[ʼndʒja] “dia”
<b>ŋ</b>	k	[ŋki.'la] “cauda de animal”
<b>ŋ</b>	g	[ŋga.'na] “enganar”
<b>m</b>	p	[ʼmpu.nɐ] “joelho”
<b>m</b>	b	[ʼmbɐ.dɐ] “lado, metade”

Fonte: Elaborado pelos autores.

A partir dos exemplos dispostos, há as seguintes possibilidades de interpretação a respeito da nasal + obstruente em posição inicial de palavra, podendo tratar-se de: (1) uma nasal silábica; (2) uma vogal nasalizada; (3) uma consoante pré-nasalizada. Discutiremos cada uma das possibilidades. Assumir, como proposto em (1) e por Bandeira et al. (2021), a consoante nasal como silábica, implicaria um custo em termos de escala, pois haveria, em angolar, um núcleo com baixa sonoridade. Ademais, vale recuperar o segundo item da Generalização de Sequenciamento Silábico (*Sonority Sequencing Generalization*): se uma língua aceita um segmento silábico com valor de sonoridade x, então todos os segmentos com valores de sonoridades maiores do que x são também potenciais núcleos silábicos (Blevins, 1995, p. 216). Portanto, se admitirmos que a consoante nasal, em angolar, é silábica, tal assunção implicaria, teoricamente, que sons com maior sonoridade, como líquidas, poderiam também

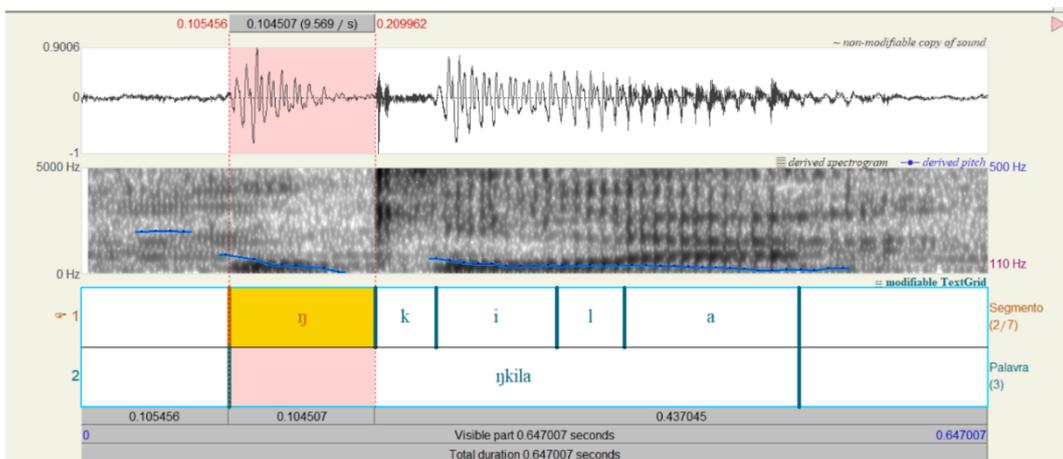
<sup>6</sup> Em virtude das limitações de nosso *corpus*, a presente análise não nos permite dizer se a sílaba CGV somente ocorre pós-lexicalmente, não sendo uma estrutura básica do léxico angolar. Por ora, os dados de variação de nosso *corpus* permitem apenas a afirmação cautelosa de que a estrutura CGV pode ser considerada uma manifestação pós-lexical, emergindo de processos fonológicos aplicados às formas lexicais. Estudos futuros deverão responder à questão referente à possibilidade de todos os glides encontrados na superfície serem oriundos de um hiato na subjacência, ou de haver, também, glides lexicais no angolar.

ocupar a posição de núcleo, porém não foram encontradas evidências para tanto. Também não é possível tratar tais segmentos como uma vogal nasalizada, dado que uma porção dos dados do nosso *corpus* de palavras com a nasal inicial não apresentou indícios para essa assunção. No caso de tais sequências, realizamos, em caráter exploratório, verificações acústicas com o *software* Praat – version 6.2.20 (Boersma; Weenink, 2022). Um exemplo da falta de pistas acústicas que suportem a presença de uma vogal no contexto referido são os itens **ngana** [ŋga.'na] “enganar” (Figura 3) e **nkila** [ŋki.'la] “cauda” (Figura 4). Nas Figuras 3 e 4, não se encontram formantes esperados em uma produção vocálica em início de palavra. Na figura 3, tal ausência vocálica é reafirmada pelo fato de que, se houvesse uma vogal nasalizada [ĩ] no início de **nkila**, seria esperado um segundo formante mais nítido, já que uma posição mais alta de F2 é atestada para a vogal alta anterior, o que não ocorre.



Fonte: Elaborado pelos autores.

Figura 3 - Produção de [ŋga.'na] “enganar” na frase-veículo **N ga fa x bathu** “Eu falo x baixo” em angolar.

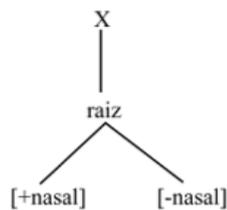


Fonte: Elaborado pelos autores.

Figura 4 - Produção de [ŋki.'la] “cauda” na frase-veículo **N ga fa x bathu** “Eu falo x baixo” em angolar.

Além de considerarmos as evidências acústicas para comprovação das consoantes pré-nasalizadas, é preciso também considerar como as pré-nasalizadas se comportam quanto à escala de sonoridade. Para tratar da questão, é necessário estabelecer antes como representar tais segmentos. Em primeiro plano, já assumimos que a nasal e a obstruente em início de palavra, em angolano, não estão em sílabas diferentes, mas em uma mesma sílaba. Contudo, se decidíssemos tratar a sequência nasal + obstruente como segmentos distintos ocupando o *onset*, haveria uma violação ao princípio, pois não se observaria uma sonoridade crescente até o pico silábico ( $1 < 0 < 3$ ), como preconiza a escala. Em contrapartida, caso se trate a sequência nasal + obstruente em início de palavra como segmento em contorno, a escala de sonoridade não seria afetada, posto que sua atuação se dá no nível segmental.

(5)



Desse modo, a violação à escala de sonoridade pode ser evitada ao tratarmos a sequência como segmento de contorno nasal (cf. Damulakis, 2010), não sendo então dois segmentos discretos, mas um segmento em contorno ocupando a posição de *onset* simples ( $[^n k, ^n g, ^m b, ^m p, ^n d, ^n t]$ ) como demonstrado em (5) (Clements; Hume; 1995). É relevante acrescentar que essa sequência, embora seja analisada como uma consoante, possui restrições fonotáticas distintas de outras consoantes. Tal segmento, por exemplo, ocorre exclusivamente no início de palavras e não participa de ataques complexos seguidos de /r/ ou /l/.

## 5.2 Síntese

No Quadro 11, são sintetizadas as estruturas de superfície que constituem o *onset* do angolano propostas por Maurer (1995), Lorenzino (1998), Bandeira (2017), Bandeira et al. (2021) e a presente análise.

Quadro 11 – Estruturas de *onset* em sílaba fonológica do angolano segundo Maurer (1995) e Lorenzino (1998), Bandeira (2017), Bandeira et al. (2021) e a análise atual.

	Maurer (1995)	Lorenzino (1998)	Bandeira (2017)	Bandeira et al. (2021)	Análise atual
V	SIM	SIM	SIM	SIM	SIM
VV	SIM	---	SIM	SIM	---
CV	SIM	SIM	SIM	SIM	SIM
CVN	---	---	SIM	SIM	SIM
CVV	SIM	SIM	SIM	SIM	---
CVVN	---	---	SIM	SIM	---
CGV	---	SIM	SIM	SIM	SIM
CGVN	---	---	SIM	SIM	SIM
N	---	---	---	SIM	---
CCV	--	--	---	SIM	SIM

CCVN	--	--	---	---	SIM
------	----	----	-----	-----	-----

Fonte: Elaborado pelos autores.

Os achados da análise atual, conforme sintetizados no Quadro 11, revelam algumas observações significativas em comparação com as propostas anteriores de Maurer (1995), Lorenzino (1998), Bandeira (2017) e Bandeira et al. (2021). Em comparação com as propostas anteriores, destacam-se os seguintes aspectos: a) confirmação de padrões anteriores; b) diferenças na consideração da nasal + obstruente. Na análise atual, reafirmamos alguns dos pontos propostos por Bandeira et al. (2021), o que sugere que esses padrões podem ser robustos e consistentes na língua. Isso fortalece a validade das conclusões anteriores e fornece mais evidências em apoio aos padrões CGV, CGVN, CCV e CCVN. Enquanto Bandeira et al. (2021) propunham que as nasais em início de palavra seguidas por obstruintes em angolar fossem classificadas como silábicas, a análise atual toma outra posição. Consideramos esses elementos como oclusivas pré-nasalizadas, tratando-as como um único segmento de contorno. Essa diferença na interpretação reflete perspectivas diferentes sobre a natureza fonológica dessas consoantes, demonstrando uma compreensão distinta a respeito da organização silábica da língua analisada.

## 6 CONSIDERAÇÕES FINAIS

O presente estudo perseguiu as seguintes metas: (i) identificar os padrões silábicos presentes na língua angolar especialmente quanto ao *onset*; (ii) observar as relações entre o ataque silábico do angolar e os Princípios da Sonoridade, incluindo a aplicação da Lei do Contato Silábico (Murray; Vennemann, 1983); (iii) comparar a análise do *corpus* com as descrições prévias de Maurer (1995), Lorenzino (1998), Bandeira (2017) e Bandeira et al. (2021). Desse modo, com relação ao primeiro objetivo, no que diz respeito às constituições silábicas no *corpus*, encontramos as seguintes possibilidades: V, CV, CGV, CCV, CVC, VC e VG.

Quanto ao segundo objetivo, a análise dos dados revelou uma organização silábica coerente com a hierarquia de sonoridade. A estrutura das sílabas segue padrões como C1V e C1C2V, respeitando a sequência de sonoridade. Essa disposição é corroborada pela preferência por *onsets* simples, compostos por uma única consoante, bem como pela ocorrência de consoantes pré-nasalizadas. Assim, as sílabas na língua são estruturadas da seguinte forma: C1V, onde C1 representa a consoante menos sonora e V representa a vogal, elemento mais sonoro, ou C1C2V, onde C1 representa uma obstruente e C2 representa uma consoante líquida, mais sonoro, do que a obstruente e menos sonora do que a vogal, e V, o pico silábico. A observância dessa hierarquia de sonoridade demonstra o cumprimento às tendências universais relacionadas à organização silábica.

De acordo com o *Princípio de Sequenciamento de Soância* (PSS), constatamos a presença de consoantes pré-nasalizadas em angolar que não podem ser prontamente classificadas como núcleos ou coda. Da mesma forma, a interpretação de tais segmentos como um *onset* complexo não é viável, uma vez que isso iria de encontro à escala de sonoridade. Nesse contexto, uma abordagem plausível é considerá-las como segmentos de contorno nasal, agrupando a sequência nasal + oclusiva como um único elemento no *onset*. Além disso, uma verificação acústica de alguns dados, de caráter exploratório, revela que a consoante nasal antecede a obstruente no início de palavras.

Por fim, quanto aos pontos de convergência e discrepância entre a análise dos dados coletados neste estudo e as descrições anteriores da estrutura silábica do angolar (Maurer, 1995; Lorenzino, 1998; Bandeira, 2017; Bandeira et al., 2021), a análise atual identificou uma estrutura silábica não abordada nas propostas de Maurer (1995) e Lorenzino (1998), como a CCVN. Além disso, confirmamos alguns padrões propostos por Bandeira (2017) e Bandeira et al. (2021), corroborando a presença desses padrões na língua. Essa confirmação reforça as conclusões anteriores e oferece mais evidências em apoio aos padrões CGV, CGVN, CCV e CCVN. Divergindo de Bandeira et al. (2021), interpretamos as nasais seguidas por obstruintes no início de palavras como oclusivas pré-nasalizadas, tratando-as como um único segmento de contorno.

Em síntese, as descobertas revelam um angolar que se alinha com princípios fonológicos universais, ao mesmo tempo em que apresenta suas peculiaridades, frutos de um contexto linguístico e social de intenso contato com o português.

## REFERÊNCIAS

- Agostinho AL. *Fonologia e método pedagógico do lung'le* [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2015.
- Agostinho AL. *Fonologia do lung'le*. München: Lincom Studies in Pidgin and Creole Linguistics. (15); 2016.
- Araujo G, Agostinho AL. Fa do Vesu, a language game of Fa d'Ambô. *PAPIA*. 2014;24(2):265-281.
- Balduino AM. *Fonologia do português de São Tomé e Príncipe* [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2022.
- Bandeira M. *Reconstrução fonológica e lexical do protocioulo do Golfo da Guiné* [tese]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2017.
- Bandeira M, Agostinho AL, Freitas S. Aspectos fonético-fonológicos do angolar moderno. *Alfa*. 2021;65:1-31.
- Bisol L. A sílaba e seus constituintes. In: Abaurre MB, organizadora. *A construção da palavra fonológica*. São Paulo: Contexto; 2013. p. 21-52.
- Blevins J. The syllable in phonological theory. In: Goldsmith J, editor. *Handbook of phonological theory*. Nova Jersey: Wiley-Blackwell Publishing; 1995. p. 206-244.
- Boersma, P., Weenink D. Praat: doing phonetics by computer. 1 jan 2022.
- Casali R. Dekereke Phonology Software Tool. 24 nov 2022. Disponível em: <https://casali.canil.ca/Dekereke/>.
- Ceita MN. *Ensaio para uma reconstrução histórico-antropológica dos angolares de S. Tomé* [trabalho de conclusão de curso]. Lisboa: Centro de Estudos Africanos, Instituto Universitário de Lisboa; 1991.
- Clements GN. The role of the sonority cycle in core syllabification. In: Kingston J, Beckman ME, editores. *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press; 1990. p. 283-333.
- Clements GN, Hume E. The internal organization of speech sounds. In: Goldsmith J, editor. *Handbook of phonological theory*. Nova Jersey: Wiley Blackwell Publishing; 1995. p. 245-306.
- Clements GN, Keyser SJ. *CV phonology: a generative theory of the syllable*. Cambridge: MIT Press; 1983.
- Collischonn G. A sílaba em português. In: Bisol L. (Org.). *Introdução a Estudos de fonologia do português brasileiro*. Porto Alegre: EDIPUCRS, 1999. p. 101-129.
- Damulakis GN. Sequenciamento de sonoridade e contorno nasal em línguas Macro-Jê. *Revista de Estudos da Linguagem*. 2010;18(1):35-49.

- Ferraz LI. A Linguistic Appraisal of Angolar. *Memoriam Antônio Jorge Dias*. 1974, v. 2: p. 177-186.
- Hagemeyer T. Initial vowel agglutination in the Gulf of Guinea creoles. In: Aboh E, Smith N, editores. *Complex processes in new languages*. Amsterdam/Filadélfia: John Benjamins Publishing Company; 2009. p. 29-50.
- INE. *Recenseamento Geral da População e da Habitação: Características Educacional da População*. São Tomé: INE; 2013.
- Ladefoged P, Maddieson I. *The sounds of the world's languages*. Londres: Blackwell; 1996.
- Lorenzino G. *The Angolar Creole Portuguese of São Tomé: its grammar and sociolinguistic history*. Munich: Lincom Europa; 1998.
- Mapmaker Interactive. *MapMaker Launch Guide [homepage]*. 2024. Disponível em: <http://mapmaker.nationalgeographic.org/>.
- Maurer P. L'apport lexical bantou en angolar. *Afrikanische Arbeitspapiere (Köln)*. 1992;29:163-174.
- Maurer P. *L'angolar. Un créole afro-portugais parlé à São Tomé*. Hamburg: Buske; 1995.
- Maurer P. Angolar. In: Michaelis SM, et al. *The survey of pidgin and creole languages*. Oxford: Oxford University Press; 2013.
- Murray RW.; Vennemann T. Sound Change and Syllable Structure in Germanic Phonology. *Language*. 1983;59(3):514-528.
- Ramos AP, Tenani LE. Análise métrica do apagamento das vogais postônicas não finais no dialeto do noroeste paulista. *Estudos Linguísticos*. 2009;38(1):21-34.
- Selkirk E. The syllable. In: Hulst HVD, Smith N, editores. *The structure of phonological representations: Part 2*. Dordrecht: Foris; 1982. p. 337-384.
- Selkirk E. *Phonology and syntax: the relation between sound and structure*. Cambridge: MIT Press; 1984.
- Zec D. The syllable. In: Lacy, P, editor. *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press; 2006. p. 161-193.



# Correlações entre percepção e produção de palavras heterotônicas do espanhol por aprendizes brasileiros

## *Correlations between perception and production of heterotonic Spanish words by Brazilian learners*

Pollianna Milan\*

Universidade Federal do Paraná, Curitiba, PR, Brasil

**Resumo:** Esta pesquisa investigou se a percepção de palavras heterotônicas do espanhol, por aprendizes brasileiros, transforma a produção das mesmas. Para isso, o estudo, que conta com uma etapa de treinamento perceptual, tem 26 aprendizes brasileiros, 17 pertencem ao grupo denominado com menos exposição acadêmica (dez que receberam treinamento e sete que não receberam treinamento) e nove do grupo com mais exposição acadêmica (quatro que receberam treinamento e cinco que não receberam treinamento). Os aprendizes ouviram heterotônicos falados por locutores hispano-falantes e, depois, produziram os mesmos heterotônicos em frases com contexto e frases-veículo. Os *Testes de Correlação de Spearman* demonstraram que há uma tendência maior de correlação entre percepção e produção no grupo com menos exposição acadêmica e que participou do treinamento perceptual.

**Palavras-chave:** Correlação entre percepção e produção de sons. Heterotônicos. Espanhol falado por brasileiros. Treinamento perceptual.

**Abstract:** This research investigated whether the perception of heterotonic Spanish words by Brazilian learners transforms their production. To this end, the research, which includes a perceptual training stage, has 26 Brazilian learners, 17 of whom belong to the group with less academic exposure (ten who have received training and seven who have not) and nine of whom belong to the group with more academic exposure (four who have received training and five who have not). The learners listened to heterotones spoken by Spanish speakers and then produced the same heterotones in sentences with context and vehicle sentences. *Spearman's correlation tests* showed that there was a greater tendency for correlation between perception and production in the group with less academic exposure and who took part in perceptual training.

**Keywords:** Correlation between perception and production of sounds. Heterotronics. Spanish spoken by Brazilians. Perceptual training.

## 1 INTRODUÇÃO

Este artigo<sup>1</sup> apresenta um recorte dos resultados de uma tese de doutorado que conduziu um estudo com diversas etapas de treinamento de percepção das palavras heterotônicas do espanhol em contraste com o português brasileiro (doravante, PB). Estamos nomeando palavras heterotônicas aquelas que, na comparação das duas

---

\* Professora Adjunta, Departamento de Literatura e Linguística (Dellin), da área de Português como Língua Estrangeira/Adicional (PLE/PLA), Universidade Federal do Paraná, Curitiba, PR, Brasil; [pmilan@ufpr.br](mailto:pmilan@ufpr.br)

<sup>1</sup> O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

línguas em questão, possuem a sílaba tônica em posição diferente, apesar da similitude da escrita e do significado. Há casos, por exemplo, que em português a palavra é paroxítona (como em *atmosfera*) e em espanhol é proparoxítona (*atmósfera*). Há, contudo, uma grande quantidade de exemplares que se enquadram na seguinte regra: ambas as palavras são paroxítonas e terminam em *-ia*. A diferença é que as paroxítonas do português apresentam uma elisão no encontro vocálico *i-a*, formando um hiato que leva a sílaba tônica, enquanto que, no espanhol, esse encontro vocálico é um ditongo e a tonicidade da sílaba, então, recai sobre a sílaba anterior, como, por exemplo, nas palavras em espanhol *a-ne-mia*, *bi-ga-mia*, *fo-bia* e *or-to-pe-dia*<sup>2</sup>.

A partir da semelhança das duas línguas em questão e também pela experiência como professora de espanhol para brasileiros, é possível constatar que essas palavras heterotônicas se tornam um desafio no ensino-aprendizado. Assim, o objetivo da pesquisa é investigar se a percepção desses heterotônicos do espanhol, por aprendizes brasileiros, transforma a produção dos mesmos. Concordamos com a premissa de que a percepção é ‘transformada’ (e não transferida) para a produção seguindo a teoria dos Sistemas Complexos (Larsen-Freeman, 2018). Ainda de acordo com essa teoria, a correlação entre percepção e produção não necessariamente ocorre de maneira direta, visto que o sistema de desenvolvimento linguístico não é linear e previsível.

Por isso, além de nos questionarmos se os heterotônicos seriam percebidos como tais pelos aprendizes brasileiros, com os testes de correlação, pretendemos investigar se esta percepção transformará a produção dos heterotônicos tanto na produção dessas palavras alvo em frases com contexto (enunciados completos) como em frases-veículo (com a palavra alvo em destaque no enunciado *Yo dije \_\_\_\_\_* ‘Eu disse \_\_\_\_\_’). Ainda, propomo-nos a investigar se essa possível correlação entre percepção e produção dos heterotônicos acontece de maneira semelhante em grupos de aprendizes brasileiros com distintos níveis de exposição ao espanhol na universidade. Ou seja, dos 26 participantes desta pesquisa, 17 pertencem ao grupo com menos exposição acadêmica (10 que receberam treinamento e 7 que não receberam treinamento) e 9, ao grupo com mais exposição acadêmica (4 que receberam treinamento e 5 que não receberam treinamento). O grupo denominado *com menos exposição acadêmica* se refere a graduandos de Letras Português/Espanhol que estavam cursando a disciplina de Língua Espanhola 1 e que, por isso, finalizaram a pesquisa com até 180 horas de exposição à língua espanhola na universidade; ao passo que o grupo denominado *com mais exposição acadêmica* estava cursando a disciplina Língua Espanhola 3, i.e., terminaram a pesquisa com até 360 horas de exposição à língua espanhola na instituição<sup>3</sup>.

A hipótese desta pesquisa é a de que, com o treinamento perceptual, haverá uma melhoria na percepção auditiva dos heterotônicos e isso resultará em transformação no modo como eles são produzidos na fala por meio da correlação entre percepção e produção. Ainda acreditamos que a percepção adequada dos heterotônicos transformará a produção dos mesmos, por meio da correlação, em ambos os grupos pesquisados (os participantes com menos e com mais exposição

<sup>2</sup> Estamos chamando de ‘elisão’ a separação silábica que ocorre no encontro vocálico no interior dessas palavras, ou seja, a separação das duas vogais. Nestes casos, cada uma ocupa um núcleo silábico distinto, constituindo, assim, duas sílabas diferentes e, por isso, um hiato.

<sup>3</sup> Conseguimos avaliar a proficiência do espanhol dos participantes considerando apenas a exposição que eles tinham de língua espanhola na universidade.

acadêmica). Para cumprir com os propósitos desta pesquisa, será feita, na seção seguinte, uma revisão teórica sobre modelos segmentais e suprasegmentais de percepção, em seguida apresentamos a Metodologia, os Resultados e Discussões e, por fim, as Considerações Finais.

## 2 MODELOS SEGMENTAIS E SUPRASEGMENTAIS DE PERCEPÇÃO

O enfoque deste estudo são as palavras heterotônicas, ou seja, estamos nos referindo ao acento que está mais diretamente relacionado aos estudos suprasegmentais. Contudo, esta pesquisa também foi inspirada nos modelos de percepção de aspectos segmentais na segunda língua (doravante, L2) (Best, 1995; Flege, 1995; Escudero, 2005) porque acreditamos, seguindo a premissa dos Sistemas Complexos, que aprendemos a nos comunicar por uma concatenação de fatores. Isso quer dizer que esses parâmetros se relacionam o tempo todo, o que resulta também em uma interação dos segmentos e dos suprasegmentos na cadeia da fala. Este fato é confirmado, por exemplo, por Lucente (2012) ao afirmar, em estudos prosódicos suprasegmentais, que o alinhamento entre segmento e suprasegmento é um dos parâmetros que se correlaciona na produção e percepção da entonação. Por isso, inicialmente apresentamos, nesta revisão de literatura, os estudos segmentais de percepção que guiaram a nossa pesquisa para, em seguida, detalharmos os suprasegmentais.

### 2.1 Modelos segmentais de percepção

Os estudos de percepção de sons, no nível segmental, foram os pioneiros em nortear, mais tarde, os estudos perceptuais dos suprasegmentos. E, desde os primeiros estudos dessa área, a relação entre produção e percepção em L2 não é algo passivo na literatura. Esta investigação começou inspirada em três modelos sobre a percepção de aspectos segmentais dos sons de uma L2, que postulam que os sons da língua materna influenciam os da língua estrangeira e que a percepção costuma vir antes da produção. São eles: o *Perceptual Assimilation Model* (doravante, PAM), de Best (1995), que pode ser traduzido como o Modelo de Assimilação Perceptual; o *Speech Learning Model* (doravante, SLM), de Flege (1995), que pode ser traduzido como o Modelo de Aprendizado de Fala – em 2021 este modelo foi revisado e chamado de *Revised Speech Learning Model* (SLM-r) (Flege; Bohn, 2021); e, mais recentemente, o *Second Language Linguistic Perception Model* (doravante, L2LP), de Escudero (2005), que pode ser traduzido como o Modelo de Percepção Linguística de uma Segunda Língua.

Best tratou mais especificamente do modelo PAM para a língua estrangeira (PAM-L2) em Best e Tylor (2007). Os autores defendem que o foco da percepção da fala está na informação sobre os eventos que produzem os sinais de fala, isto é, na fonologia articulatória, e, a partir dessas questões articulatórias, o ouvinte faz uma assimilação variável<sup>4</sup> dos sons da segunda língua aos de sua primeira língua (doravante, L1). Os autores lembram ainda que a percepção de um ouvinte ‘ingênuo’, ou seja, que

<sup>4</sup> O conceito de assimilação ‘variável’ no modelo PAM-L2 se refere ao fato de que é possível perceber as diferenças linguísticas tanto em relação ao caráter contrastivo (dois léxicos diferentes, com seqüências sonoras muito parecidas) como em aspectos variáveis, mas não distintivos. Além disso, a assimilação variável leva em consideração também a abordagem multimodal da fala, em que outros elementos, como os acústicos e os visuais, atuam na categorização do sistema fônico.

fala apenas uma língua, é diferente daquele que tem experiência com outras línguas além da materna. O modelo PAM-L2 prevê, por exemplo, que quanto mais próximos estiverem os sons da L2 com os da L1, mais facilmente eles serão assimilados como uma categoria da L1 e, assim, mais difícil será o aprendizado. Como este presente estudo tem como alvo os heterotônicos do espanhol e esses são muito próximos às palavras do PB, inclusive na questão da percepção do som, isso pode ser um dificultador, segundo o PAM-L2.

O modelo SLM proposto por Flege (1995) segue uma linha semelhante ao PAM-L2 no que concerne à ideia de que a similitude dos sons da língua materna (L1) influencia ou não o desenvolvimento da segunda língua (L2) em um mecanismo de interação entre ambas as línguas por meio da assimilação de categorias fonéticas e da dissimilação de categorias. Isso quer dizer que a probabilidade de perceber a diferença de um som da L1 para um som da L2 aumenta quando diminui a semelhança entre eles. Então, um som muito parecido da L2 será produzido como o da L1, porque não há a criação de uma nova categoria para ele no espaço fonológico. O contrário acontece com sons diferentes, que recebem uma categoria nova. O autor ressalta ainda que o modelo não afirma, no entanto, que todos os erros de produção de L2 são perceptivamente motivados. Mesmo assim, o SLM é um dos modelos que acredita que a percepção antecede a produção. Essa premissa não é mantida na versão revisada (SLM-r). Flege e Bohn (2021) defendem, nesta releitura do modelo, que percepção e produção se desenvolvem concomitantemente. Ainda, o modelo original considerava a criação de novas categorias fonéticas por aprendizes de nível avançado na L2. Já no modelo revisado, isso foi considerado também para aprendizes iniciantes na língua alvo.

O modelo L2LP prega que o estudante de uma língua estrangeira usa a sua gramática da percepção, ou seja, de sua L1, quando começa a adquirir essa L2. Escudero (2005, p. 138) chama isso de *Full Copying Hypothesis*. Segundo a autora, essa hipótese só pode ser testada, se o cruzamento linguístico ou a percepção não nativa da L2 é comparada à percepção de L2 de iniciantes absolutos nessa língua estrangeira. Por exemplo, a percepção das vogais dos falantes britânicos do Sul por ouvintes espanhóis só deve ser comparada à percepção de L2 de aprendizes espanhóis iniciantes desse inglês. O objetivo do modelo é, segundo Colantoni et al. (2015), prever uma sequência de desenvolvimento para cada ouvinte 'ingênuo' (chamado pelas autoras de *naïve*), aquele do primeiro contato com a língua e, assim, testar se o desenvolvimento previsível é devido ao fato observado em aprendizes de L2 iniciantes, intermediários ou avançados com uma L1 compartilhada. Assim, dados longitudinais de falantes iniciais que gradualmente ganham experiência acadêmica da L2 é o mais desejável. O modelo L2LP é semelhante ao PAM/PAM-L2 e ao SLM-r, ao usar informações acústicas para prever os padrões de categorização de cruzamentos linguísticos. Além disso, também postula que o falante de L2 tem a gramática perceptual de sua L1 que poderá influenciar na percepção dessa segunda língua.

Esses estudos iniciaram as indagações sobre a percepção preceder (ou não) a produção, questão essa que continua a ser indagada na área. Llisterri (1995) e Rochet (1996) afirmam existir motivos extra fonológicos para que a percepção preceda a produção, como a pressão social que exige uma produção adequada e as consequências comunicativas de uma produção inadequada. Na linha teórica dos Sistemas Complexos, Caspi e Lowie (2013) relatam que a recepção do vocabulário de uma L2 para a produção do mesmo está longe de ser linear, uma característica dessa teoria. Em

um estudo de caso detalhado sobre o desenvolvimento do conhecimento de vocabulário de inglês como L2, por um período de 36 semanas de exposição intensa (fala e escrita acadêmica) a essa L2 por uma aprendiz brasileira, os resultados mostraram que os níveis receptivo e produtivo do conhecimento interagem de forma complexa ao longo do tempo, às vezes competindo e às vezes apoiando-se mutuamente. Nesse estudo, o participante demonstrou que a receptividade de itens do vocabulário é transformada para a produção, mesmo em um ambiente de grande variabilidade. Isso quer dizer que a receptividade do vocabulário antecedeu a produção, indo ao encontro do que postula, por exemplo, o SLM de Flege.

Apesar de muitos estudos já comprovarem que a percepção precede a produção, há outros (Flege; Bohn, 2001; Altmann, 2006; Kijak, 2009) que demonstram o contrário ou questionam se realmente essa é uma relação de causa e efeito, ou ainda, se uma interação (entre percepção e produção) pode sofrer alterações ao longo do desenvolvimento de uma língua. Como veremos a seguir, nos modelos de percepção suprasegmental, não necessariamente perceber significa produzir adequadamente ou vice-versa. Trataremos sobre esse assunto na próxima seção.

## 2.2 Modelos suprasegmentais de percepção

Há duas linhas de pesquisa que adotaram modelos de percepção suprasegmental que nos interessam sobremaneira, pois tratam também do acento, que é o foco desta investigação. A primeira delas é a *Stress Deafness Model* (doravante, SDM), ou Modelo de ‘Surdez’ Acentual, que teve origem nos estudos de Dupoux et al. (1997, 2008). A segunda linha chama-se *Stress Typology Model* (doravante, STM), ou Modelo de Tipologia do Acento, que se baseia, principalmente, na pesquisa de Altmann (2006). A seguir, ilustraremos do que tratam cada um dos dois modelos.

### 2.2.1 Os modelos SDM e STM

O modelo SDM trabalha com o conceito de ‘surdez’ acentual que designa, segundo Peperkamp e Dupoux (2002, p. 204), o efeito de os ouvintes terem dificuldades perceptuais na discriminação de palavras inexistentes no seu idioma que formam um par mínimo de certos contrastes fonológicos não-nativos, seja segmental ou suprasegmental. Isso quer dizer que esses ouvintes falham na percepção auditiva desses contrastes. A palavra ‘surdez’ foi mantida entre aspas tal como tem sido usada pelos criadores dessa corrente teórica. Esse modelo surgiu em estudos de percepção acentual no cruzamento das línguas espanhola e francesa. O primeiro deles foi desenvolvido a partir de quatro experimentos em que Dupoux et al. (1997) observaram a ‘surdez’ acentual dos franceses com relação a palavras consideradas pares mínimos do espanhol (que se distinguiam apenas por uma diferença acentual). Essa possível ‘surdez’ dos franceses com relação ao acento deve-se ao fato de que a língua francesa não usa esse parâmetro para distinguir duas palavras iguais ou semelhantes, resultando em uma inexistência da representação mental fonológica do acento nesses indivíduos. Para Dupoux e Peperkamp (2002, p. 169, tradução nossa), “durante o primeiro ano de vida, crianças adquirem muitas propriedades fonológicas de sua língua nativa e perdem

a sensibilidade dos contrastes fonológicos que não lhe são pertinentes.”<sup>5</sup> Essa é uma das hipóteses defendida por Dupoux et al. (1997), e por Dupoux e colaboradores em trabalhos posteriores, para explicar a ‘surdez’ francesa.

Inspirados nas propostas desenvolvidas pelo modelo SDM, Vogel (2000), Altmann e Vogel (2002) e Altmann (2006) desenvolveram outros experimentos de cruzamentos linguísticos que deram origem ao STM. Esse modelo é bastante semelhante ao SDM, com a diferença de investigar línguas de acento imprevisível e línguas sem acento. Assim, o STM classifica as línguas de acordo com o acento (igualmente ao modelo SDM), mas de uma maneira hierárquica. Como só tivemos acesso ao estudo de Altmann (2006), é nele que nos asseguramos para descrever o STM.

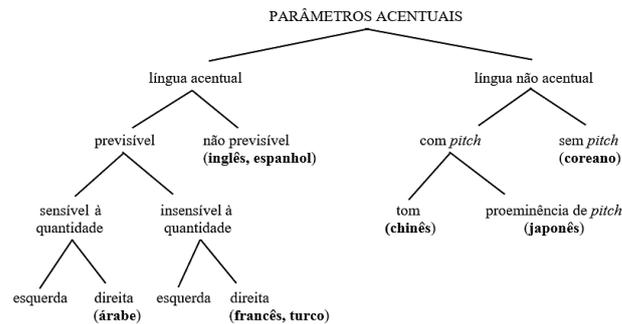
Assim como no SDM, o que norteia o modelo STM é a premissa de que os parâmetros acentuais da L1 parecem influenciar no modo como os aprendizes de uma L2 percebem e acentuam as palavras dessa língua a ser desenvolvida. Uma das perguntas que Altmann (2006) se faz é quão corretamente falantes de idiomas tipologicamente diferentes podem localizar o acento do inglês como L2. Em um estudo desenhado para investigar a percepção acentual de L2, Altmann e Vogel (2002 apud Altmann, 2006) examinaram a habilidade de aprendizes de inglês como L2, de diferentes línguas maternas, em localizar o acento primário em palavras inexistentes, mas possíveis do inglês. Os resultados desse primeiro experimento mostraram que falantes de língua sem acento (coreanos e chineses), com ou sem tom, foram muito bem, assemelhando-se aos resultados dos falantes de inglês como L1 (grupo controle). Falantes de língua materna que têm o acento fonologicamente previsível (como o turco e o tailandês) foram piores do que todos os outros. Assim, Altmann e Vogel (2002 apud Altmann, 2006) concluíram que a tipologia da língua parece influenciar na percepção do acento.

Altmann (2006) explica que o modelo é chamado de tipológico porque seus autores classificam e agrupam os idiomas com base em suas propriedades de acento de palavras (ou parâmetros) de maneira hierárquica. Altmann (2006, p. 26, tradução nossa) recorda ainda que Vogel (2000) “postula uma tipologia de sistema de acento que inclui vários tipos diferentes de línguas que especificamente consideram configurações de L2 e não apenas a perceptibilidade do acento em geral.”<sup>6</sup> Dessa maneira, Vogel (2000 apud Altmann, 2006) traçou um gráfico em que demonstra como essas hierarquias funcionam. Apresentamos o mesmo na Figura 1, adaptado para o estudo de Altmann (2006) que inclui, além da tipologia acentual, algumas línguas abordadas por ele em sua tese de doutorado.

---

<sup>5</sup> “During the first year of life, infants acquire many phonological properties of their native language and lose their sensitivity to phonological contrasts that are not pertinent.” (Dupoux; Peperkamp, 2002, p. 169).

<sup>6</sup> “Posit a stress system typology that includes various different types of languages which specifically considers L2 settings and not merely the perceptibility of stress in general.” (Altmann, 2006, p. 26).



Fonte: Vogel (2000 apud Altmann, 2006, p. 31).

Figura 1 – Gráfico hierárquico da tipologia das línguas.

Segundo as descrições da Figura 1, dadas por Altmann (2006), em línguas acentuais, o acento pode ser previsível ou não. Para línguas de acento não previsível (como o inglês, o espanhol e também o PB), o acento precisa ser especificado lexicalmente e necessita codificação na representação lexical das palavras. Se o acento é previsível, geralmente não há requisito lexical. Também depende se a atribuição regular do acento requer conhecimento sobre o peso da sílaba, por isso as línguas são classificadas como sensíveis à quantidade (peso silábico), como o árabe, ou insensíveis, como o francês e o turco. O STM leva em consideração as línguas não acentuais (o que o SDM não faz). Segundo a Figura 1, essas línguas podem usar o *pitch* para marcar a sílaba acentuada, a partir do tom (como o chinês), ou da proeminência de *pitch* (como o japonês). Há, ainda, as línguas não acentuadas que não fazem uso de *pitch*, como o coreano. Segundo Altmann (2006, p. 32, tradução nossa), “a hierarquia tipológica apresentada aqui prevê diferentes graus de dificuldade em relação ao desenvolvimento de acento primário em uma L2 por falantes de diferentes L1s<sup>7</sup>.” Por isso, o melhor desempenho para o acento de uma L2 seria esperado por falantes de língua materna sem acento, já que não haveria configurações de parâmetro acentual de L1 positivas que poderiam interferir nos ajustes de uma L2. Por outro lado, os falantes de L1 com acento previsível poderiam apresentar maiores dificuldades de acordo com essa tipologia.

O estudo de Altmann (2006) contou com aprendizes avançados de inglês como L2 de sete distintos grupos de L1 (10 de cada língua): árabe, chinês, francês, japonês, coreano, espanhol e turco (suas tipologias estão distribuídas na Figura 1). Falantes de inglês como língua materna participaram do experimento para servir de referência aos resultados dos grupos linguísticos a serem avaliados. Altmann (2006, p. 39) escolheu as línguas turca e árabe para comporem o conjunto de línguas de acento previsível porque elas sequer foram citadas no modelo SDM. A hipótese do estudo é a de que, se todos os participantes adquirirem bem o acento do inglês, a L1 não influencia no desenvolvimento acentual da L2. Se isso não acontecer, poderá haver variações individuais ou de grupos linguísticos tipológicos definidos pela L1.

O experimento era composto de palavras inventadas (e possíveis no inglês) com duas, três ou quatro sílabas com a sequência CV<sup>8</sup>. Havia teste de percepção e

<sup>7</sup> “The hierarchical typology presented here predicts different degrees of difficulty regarding the acquisition of primary stress in a certain L2 by speakers of different L1s.” (Altmann, 2006, p. 32).

<sup>8</sup> CV quer dizer ‘consoante + vogal.’

produção de acento e, para cada um deles, foi usado um conjunto distinto de palavras. No teste de percepção, foram julgadas 125 palavras que apareciam, cada uma, na tela do computador, separadas silabicamente. Os participantes ouviam a palavra e deviam clicar na sílaba que lhes parecia ser a acentuada. Os resultados para os testes de percepção, pelo menos para as línguas de acento previsível, foram os mesmos dos estudos de Dupoux et al. (1997).

Falantes de árabe, turco e francês tiveram acertos piores. Por outro lado, houve uma tendência clara de acertos tanto quanto os dos falantes de inglês como L1 para os falantes de japonês, coreano, chinês e espanhol (línguas sem acento ou com acento imprevisível). Assim, os resultados foram congruentes com o modelo STM. Os falantes de línguas com acento previsível demonstraram desempenho relativamente ruim em comparação aos outros. Os falantes de espanhol, a língua mais parecida com o inglês nesse estudo, por apresentar acento fonologicamente imprevisível, tiveram um desempenho semelhante ao dos falantes de inglês como L1. Além disso, os falantes de idiomas sem acento lexical também tiveram um desempenho extremamente bom. Assim, o estudo aqui proposto, na linha do que registra a literatura, pode apresentar resultados positivos seguindo o que prevê o modelo STM, de que línguas com acento variado têm mais sucesso na percepção acentual do que línguas com acento fixo. Como o PB e o espanhol têm acento variado, isso poderá facilitar a percepção dos heterotônicos do espanhol pelos brasileiros aprendizes dessa língua.

No teste de produção do estudo de Altmann, os mesmos participantes receberam uma lista de palavras novas (inexistentes, mas possíveis no inglês) e deveriam ler em voz alta cada palavra da lista, que continha 46 novos estímulos que não ocorreram no teste de percepção. Os resultados surpreenderam Altmann (2006). O grupo controle, dos falantes maternos de inglês, preferiu acentuar essas palavras desconhecidas no padrão paroxítono. Isso mostrou-se ainda mais forte em palavras de duas sílabas no padrão CVCV. Curiosamente, os franceses, que no modelo SDM foram considerados ‘surdos’ acentualmente e que no modelo STM demonstraram resultados piores para o teste de percepção, foram tão bem quanto os ingleses no teste de produção. Os turcos, espanhóis e árabes ficaram indecisos. Os chineses, coreanos e japoneses claramente preferiram acentuar a sílaba final e não a penúltima sílaba, como os falantes de inglês (L1). Segundo Altmann (2006, p. 129), um padrão geral de erro que emergiu entre os falantes de L2 foi o posicionamento incorreto do acento na sílaba final, que é precisamente a sílaba que os falantes nativos de inglês evitavam acentuar. Curiosamente, porém, os falantes de línguas de acento previsível foram melhores na produção do que na percepção. Isso quer dizer que os falantes de línguas que mostraram uma pior taxa na percepção do acento apresentaram desempenho bom na produção. Além disso, os grupos de falantes de segunda língua que tiveram as melhores pontuações de percepção produziram palavras com os padrões de acento menos parecidos aos produzidos pelos ingleses. Em outras palavras, essas configurações acentuais positivas (especificamente, acento previsível e na fronteira direita da palavra) do francês e do turco parecem gradualmente impedir a capacidade de perceber o acento, embora pareçam ajudar na produção do mesmo. Esse cenário indica que os falantes com experiência na produção de acento fonológico na sua primeira língua, seja esse acento previsível ou imprevisível, parecem ter uma vantagem sobre as línguas não acentuais.

Assim, pode-se induzir que deficiências nas habilidades de percepção ou ‘surdez’ acentual não significam que os falantes que apresentam essas características

não sejam capazes de perceber os contrastes de acento. Altmann (2006) finaliza o experimento com uma questão que nos chama bastante a atenção: a falta de correlação entre percepção e produção. Para ele, “a suposição comum de que boa capacidade perceptiva é um pré-requisito para o sucesso na produção [como indicado no modelo SLM de Flege (1995) ou no PAM de Best (1995)] parece ser infundada, pelo menos no que diz respeito ao acento.” (Altmann, 2006, p. 141-142, tradução nossa)<sup>9</sup>.

A maioria dos estudos fonológicos, como o de Altmann (2006), prevê a investigação dos efeitos de uma segunda língua na língua materna a partir do inglês como L2. Kijak (2009) faz um movimento contrário, analisando o efeito do polonês como L2 em distintas L1s, a partir das suposições dos modelos SDM e STM até então aqui apresentados. O polonês é uma língua com atribuição regular de acento, pois ele é fixado na penúltima sílaba. A autora usou essa língua, que é diferente do inglês, para ser o padrão de segunda língua a ser percebida e produzida por falantes das seguintes L1s: russo, tcheco, alemão, francês, inglês, espanhol, italiano e chinês. O objetivo de Kijak (2009) foi investigar como as diferenças linguísticas nos tipos de sistemas métricos de falantes de uma variedade de L1 afetam o desenvolvimento do acento de polonês como L2. Os resultados de Kijak (2009) mostram que não basta olhar apenas para as regras acentuais da L1 para explicar o desenvolvimento acentual da L2. Para ela, a função que o acento desempenha em L1 também pode ser crucial. O acento imprevisível e lexicamente codificado, em russo, por exemplo, tem uma importante função contrastiva e seus falantes maternos foram muito bem no teste de percepção acentual aplicado pela autora. No entanto, em outras línguas como inglês, alemão, espanhol e italiano, o acento tem apenas uma função parcialmente contrastante, pois também é parcialmente previsível, exigindo um menor grau de codificação. Kijak esperava, assim, que os falantes dessas línguas também fossem razoavelmente bem na percepção do acento, de acordo com os modelos STM e SDM. No entanto, os falantes de inglês, alemão, espanhol e italiano, apesar dos padrões acentuais semelhantes, diferiram muito em seus resultados de percepção acentual, com os ingleses apresentando desempenhos significativamente piores do que os outros. Uma explicação possível para isso, segundo Kijak, é que:

Para os falantes de inglês, as informações acentuais não são tão cruciais no processo de reconhecimento de palavras em L1 quanto nos falantes das outras línguas. Nós sugerimos que isso também tem consequências para a sua percepção acentual: quanto menos informações de acento ‘usadas’ no processamento da fala em L1, mais fraca a capacidade de percepção de acento de seus falantes<sup>10</sup> (Kijak, 2009, p. 319, aspas simples da autora, tradução nossa).

Além disso, a provável ‘surdez’ esperada para os tchecos, por exemplo, que têm um acento extremamente previsível, não foi encontrada nos estudos de Kijak, mostrando que um acento de L1 extremamente regular não conduz necessariamente à ‘surdez’ acentual. Nos resultados de produção, Kijak (2009) encontrou uma clara

<sup>9</sup> “The common assumption that good perceptual ability is a prerequisite for success in production (as indicated in the model SLM by Flege (1987) or in PAM by Best (1995)) thus seems to be unfounded, at least with regard to stress.” (Altmann, 2006, p. 141-142).

<sup>10</sup> “For English speakers the stress information is not as crucial in the word recognition process in L1 as it is for speakers of the other languages. We suggested that this also has consequences for their stress perception: the less ‘used’ stress information in speech processing in L1, the poorer its speakers’ stress perception ability.” (Kijak, 2009, p. 319, aspas simples da autora).

divisão entre falantes de alemão, espanhol, italiano, russo e inglês, por um lado, e falantes de francês, tcheco e chinês por outro lado. O primeiro grupo teve um desempenho consideravelmente melhor do que o segundo. As taxas de sucesso observadas para a produção, porém, não foram refletidas diretamente nas taxas de sucesso da percepção. No estudo de Kijak (2009), tchecos foram muito bem na percepção do acento polonês, mas eles foram mal na produção desse acento. Ao contrário, falantes de inglês foram mal na percepção do acento polonês, mas foram muito melhores na produção do mesmo. Ou seja, no estudo de Kijak, a relação percepção e produção variou muito e nem sempre foi direta.

Os estudos apresentados até aqui, de Dupoux et al., Altmann e Kijak, mostram que ainda há muito a ser pesquisado, apesar de esses já apresentarem discussões e resultados interessantes sobre a percepção e a produção de acento de uma segunda língua. Seguimos as afirmações de Altmann (2011), de que precisamos nos questionar (e sobretudo pesquisar), se a percepção do acento é uma habilidade/aptidão ou se se deve a um conjunto de fatores ainda não totalmente explícitos. Nesse ponto, conforme Altmann (2011), pode-se afirmar que os aprendizes de L2 são afetados pelas propriedades prosódicas e fonéticas de sua língua materna, mas podem ser capazes de empregar outros aspectos e mecanismos gramaticais internos do idioma para compensar o desempenho da L2.

Até aqui foi possível demonstrar que tanto os modelos SDM e STM, assim como os achados de Kijak, podem fornecer pistas que nos ajudarão a compreender as facilidades e dificuldades a serem apresentadas pelos aprendizes brasileiros com relação aos heterotônicos do espanhol. Na próxima seção, descrevemos a metodologia deste estudo e, na seção que trata dos resultados e discussões, voltamos a nos guiar à luz desses experimentos.

### 3 METODOLOGIA

Esta pesquisa objetiva investigar como ocorre a relação entre percepção e produção de palavras heterotônicas do espanhol por aprendizes brasileiros. Também tem como propósito analisar se a percepção afeta a produção dessas palavras para cada um dos grupos investigados. Para isso, contou com um corpus de 115 palavras heterotônicas do espanhol em comparação com o PB. Esses 115 heterotônicos foram divididos entre os testes (pré-teste, pós-teste, retenção e generalização) e as duas sessões de treinamento. Para os testes, também usamos 30 palavras distratoras, selecionadas entre aquelas que costumam ter alta frequência de uso no espanhol. Todos os testes compreenderam a produção das palavras alvo como a percepção das mesmas.

O corpus para os testes de percepção foi gravado por oito locutores falantes de espanhol como L1 que se encontravam no Brasil no momento desta pesquisa: quatro deles eram mexicanos e suas locuções foram usadas nos testes em geral (pré-teste, pós-teste e teste de retenção) e nas duas sessões de treinamento. Outros quatro locutores (dois hondurenhos e dois cubanos) foram gravados e suas locuções foram usadas no teste de generalização de percepção. As gravações com os locutores consistiam na leitura, em voz alta, de frases-veículo que estavam na tela do computador, como *Yo dije atmosfera* 'Eu disse atmosfera'. Depois da gravação com os locutores, os testes de percepção foram montados no programa TP (Raubert et al., 2013) e foram validados por outros quatro falantes de espanhol como língua materna,

distintos dos locutores desta pesquisa, para avaliarmos a qualidade dos estímulos e dos testes antes de os mesmos serem aplicados.

Os aprendizes que participaram deste estudo, todos falantes de PB como L1 e que estudavam língua espanhola como L2, eram graduandos do curso de Licenciatura em Letras Espanhol da Universidade Federal do Paraná. Dos 26 participantes desta pesquisa, 17 pertencem ao grupo com menos exposição acadêmica (dez que receberam treinamento e sete que não receberam treinamento) e nove ao grupo com mais exposição acadêmica (quatro que receberam treinamento e cinco que não receberam treinamento). O grupo denominado *com menos exposição acadêmica* se refere a graduandos que estavam cursando a disciplina de Língua Espanhola 1 e que, por isso, finalizaram a pesquisa com até 180 horas de exposição à língua espanhola na universidade; ao passo que o grupo denominado *com mais exposição acadêmica* estava cursando a disciplina Língua Espanhola 3, i.e., terminaram a pesquisa com até 360 horas de exposição à língua espanhola na instituição. Como esta investigação está focada no acento do espanhol, averiguamos se os participantes tiveram aulas de acentuação sobre o espanhol antes e/ou durante a coleta de dados. Todos os aprendizes tiveram aula expositiva e exercícios sobre acentuação do espanhol, o que implica que deveriam saber pronunciar os heterotônicos. Além disso, os participantes com mais exposição acadêmica estudaram, em uma disciplina da graduação de língua espanhola, 58 heterotônicos<sup>11</sup> que apareceram nos testes, ou seja, esse grupo tinha mais familiaridade com os heterotônicos.

Nosso estudo de treinamento perceptual seguiu o padrão de testes que é comum a este tipo de pesquisa. Antes de iniciar as sessões de treinamento propriamente ditas, todos os informantes fizeram o pré-teste de produção (que avaliou a pronúncia dos heterotônicos) e, em seguida, o pré-teste de percepção (que avaliou a audição dos heterotônicos). Em todos os testes de produção, havia a tarefa de leitura de frases com contexto (que continham heterotônicos) e frases-veículo. Os pré-testes serviram para avaliarmos se os participantes conheciam, produziam e escutavam adequadamente os heterotônicos antes de nossa intervenção. Depois, uma parte dos informantes (14 deles) fez as duas sessões de treinamento perceptual para serem comparados aos outros 12 que não treinaram perceptualmente. As duas sessões de treinamento aconteceram nas três semanas seguintes aos pré-testes, conforme disponibilidade das turmas. O próximo passo, que também ocorreu nas três semanas seguintes após o treinamento, foi replicar os mesmos testes (os pré-testes), por isso são chamados de pós-testes (tanto de produção como de percepção). Juntamente com os pós-testes, foram aplicados os testes de generalização de produção e de percepção a todos os participantes (em que apareceram novos heterotônicos ainda não vistos nos outros testes e sessões de treinamento e novos locutores no teste de percepção). Por fim, entre 42 e 58 dias após a última sessão de treinamento perceptual, foram aplicados os testes de retenção de produção e percepção (iguais aos pré-testes) para averiguar se o aprendizado que pode ter ocorrido nas sessões de treinamento foi mantido a longo prazo.

Nos testes de produção, cada informante foi conduzido individualmente a uma cabine com tratamento acústico para ler as frases com contexto e frases-veículo que

---

<sup>11</sup> O docente responsável pela disciplina nos forneceu todo o material didático usado nas aulas de língua espanhola e, a partir desse material, foi possível contabilizar quantos heterotônicos os estudantes haviam visto em aula e que também estavam presentes em nosso estudo.

continham heterotônicos. No dia da aplicação dos testes de percepção, a turma inteira era conduzida ao laboratório de informática para a realização do teste e cada participante respondia o teste em um computador individual. Os testes de percepção (pré-teste, pós-teste e teste de retenção eram iguais) continham os mesmos 40 heterotônicos e os mesmos 20 distratores falados nos testes de produção. As duas sessões de treinamento (aplicadas após os pré-testes) aconteceram em dias distintos e com apenas uma parte dos participantes. Foram usados 56 heterotônicos não contemplados em nenhum dos testes de produção e de percepção e divididos entre as duas sessões: 29 para a primeira e 27 para a segunda. Como se trata de treinamento, não utilizamos palavras distratoras. As duas sessões também foram montadas no programa TP e, diferentemente dos testes perceptuais, as sessões de treinamento forneceram uma resposta (feedback imediato) a cada escolha da sílaba tônica pelo participante. Os testes de generalização (que continham 19 novos heterotônicos e 10 novos distratores) foram aplicados juntamente com os pós-testes.

A primeira parte do estudo (a do treinamento perceptual) comprovou que houve uma certa facilidade na percepção adequada dos heterotônicos por ambos os grupos (Milan, 2019) e que o treinamento perceptual foi positivo na produção. Neste artigo, questionamos se essa percepção foi transformada na produção adequada por correlação. Assim, focamos a análise dos dados apenas na possível correlação que possa ou não existir entre o que era percebido como heterotônico, nos testes de percepção, e como este era produzido nos testes de produção, em cada uma das fases (pré-testes, pós-testes, testes de generalização e de retenção). Como os dados apresentaram um comportamento não paramétrico, optamos pelo teste de *Correlação de Spearman* para averiguar se os grupos fizeram associação entre os acertos da produção e os acertos da percepção (e vice-versa). Nesse caso, além de observarmos a significância de  $p \leq 0,05$ , retratamos a avaliação de força e a direção dessa correlação, conforme sugere Martins (2011). Informaremos o valor do teste apenas para os casos em que este for significativo.

## 4 RESULTADOS E DISCUSSÕES

Os resultados desta pesquisa serão apresentados da seguinte maneira: em cada subseção, serão reportados os dados do *Teste de Correlação de Spearman* para os quatro grupos desta pesquisa e também conforme a etapa do estudo, i.e., se pré-teste, pós-teste, teste de generalização ou de retenção. Serão apresentados os dados sempre na mesma ordem: primeiro sobre o grupo com treinamento e menos exposição acadêmica, depois sobre o grupo com treinamento e mais exposição acadêmica, seguido do grupo sem treinamento e com menos exposição acadêmica e, por fim, o grupo sem treinamento e com mais exposição acadêmica. Iniciamos com os resultados do pré-teste.

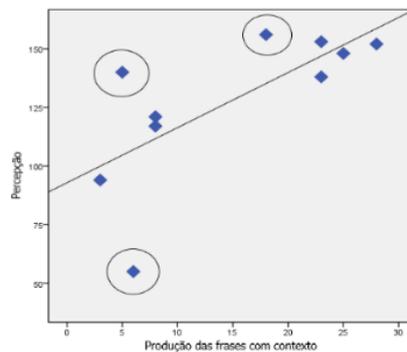
### 4.1 Resultado da correlação entre pré-testes de produção e de percepção

O grupo com menos exposição acadêmica e que participou do treinamento perceptual, segundo o *Teste de Correlação de Spearman*<sup>12</sup>, apresentou uma correlação positiva, forte e significativa entre o que produziu dos heterotônicos nas frases com contexto e o que percebeu. Isso quer dizer que produzir adequadamente os

<sup>12</sup> Os valores do *Teste de Correlação de Spearman* são  $r=0,66$ ,  $p=0,038$ .

heterotônicos inseridos nas frases com contexto esteve associado ao fato de perceber adequadamente esses mesmos heterotônicos em testes auditivos (e vice-versa). A produção esperada dos heterotônicos nas frases-veículo, porém, não esteve associada à percepção dos mesmos para esse grupo. Acreditávamos que a correlação aconteceria, principalmente, entre percepção e produção das frases-veículo e não das frases com contexto, justamente porque, nas primeiras, os heterotônicos estão em destaque. Contudo, nos pré-testes, ocorreu o contrário, o que talvez seja justificado, conforme Caspi e Lowie (2013), pelo fato de os níveis receptivo e produtivo do conhecimento interagirem de forma complexa ao longo do tempo. Ou seja, a correlação não ocorreu onde esperávamos, porque o desenvolvimento do acento não é linear e previsível.

A Figura 2 mostra a tendência dessa associação positiva entre perceber e produzir as frases com contexto desse grupo. A forte relação entre as duas variáveis apontou que, dos dez participantes do grupo com treinamento e que tinham menos exposição acadêmica, representados por um losango azul no gráfico, sete deles estiveram próximos da linha reta, ou seja, de uma tendência linear entre produzir e perceber, e apenas três deles (circulados) se afastaram dessa tendência.



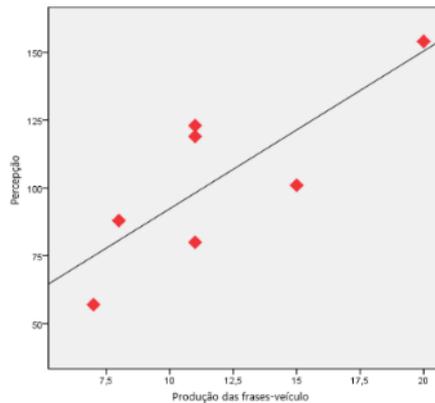
Fonte: Elaboração própria.

Figura 2 – Correlação entre produção das frases com contexto e percepção do grupo com treinamento e menos exposição acadêmica. *Scatterplot* rodado no programa *SPSS*.

O grupo com treinamento e com mais exposição acadêmica não apresentou correlação entre o que produziu e o que percebeu nos pré-testes. O grupo sem treinamento e com menos exposição acadêmica, segundo o *Teste de Correlação de Spearman*<sup>13</sup>, apresentou uma correlação positiva, forte e significativa entre o que produziu dos heterotônicos nas frases-veículo e o que percebeu. Assim, produzir adequadamente os heterotônicos inseridos nas frases-veículo esteve associado ao fato de perceber adequadamente essas palavras em testes auditivos (e vice-versa). A produção esperada dos heterotônicos nas frases com contexto para esse grupo, porém, não esteve associada à percepção correta dos mesmos. Importante notar que ocorreu justamente o contrário com o grupo de mesmo nível, mas que treinou.

A Figura 3 mostra essa relação e, diferentemente do grupo com treinamento e menos exposição acadêmica, os sete aprendizes, representados por um losango vermelho, mantiveram uma distância razoavelmente semelhante da tendência linear e apenas um deles esteve mais próximo da linha de tendência de correlação do que os outros.

<sup>13</sup> Os valores do *Teste de Correlação de Spearman* são  $r=0,75$ ,  $p=0,057$ .



Fonte: Elaboração própria.

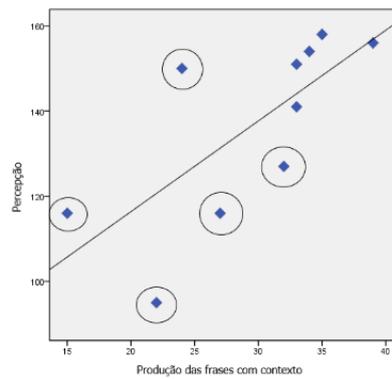
Figura 3 – Correlação entre produção das frases-veículo e percepção do grupo sem treinamento e com menos exposição acadêmica. *Scatterplot* rodado no programa *SPSS*.

O grupo sem treinamento e com mais exposição acadêmica também não apresentou correlação entre o que produziu e o que percebeu nos pré-testes. Como os dois grupos com mais experiência acadêmica em espanhol, com e sem treinamento, não apresentaram correlação, talvez isso possa indicar que ter uma exposição maior à língua alvo na universidade (ou seja, mais horas em disciplinas cursadas) faz com que esses grupos não utilizem a correlação entre percepção e produção (e vice-versa) na hora de realizar os testes. Vamos aos resultados do pós-teste.

#### 4.2 Resultados da correlação entre pós-testes de produção e de percepção

O grupo com treinamento e menos exposição acadêmica, segundo o *Teste de Correlação de Spearman*<sup>14</sup>, apresentou uma correlação positiva, forte e significativa entre o que produziu dos heterotônicos nas frases com contexto e o que percebeu. Isso quer dizer que produzir adequadamente os heterotônicos inseridos nas frases com contexto esteve associado ao fato de perceber adequadamente esses mesmos heterotônicos em testes auditivos (e vice-versa). A Figura 4 mostra a tendência dessa associação positiva entre perceber e produzir as frases com contexto. A forte relação entre as duas variáveis apontou que, dos dez participantes deste grupo, representados por um losango azul escuro no gráfico, cinco aprendizes ficaram bastante próximos da linha de tendência de correlação e cinco deles, que estão circulos na Figura 4, mantiveram uma distância maior da linha de tendência. Segundo os Sistemas Dinâmicos (Larsen-Freeman, 2018), essa variação em um mesmo grupo é comum, dado o fato de que os aprendizes costumam desenvolver uma L2 em tempos diferentes, mesmo que pertençam a um nível semelhante de proficiência.

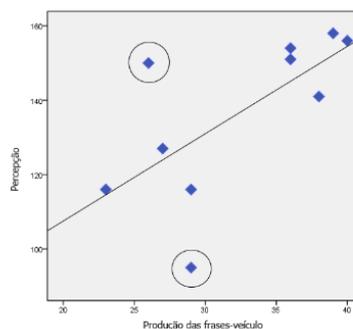
<sup>14</sup> Os valores do *Teste de Correlação de Spearman* são  $r=0,88$ ,  $p=0,001$ .



Fonte: Elaboração própria.

Figura 4 – Correlação entre produção das frases com contexto e percepção do grupo com treinamento e menos exposição acadêmica. *Scatterplot* rodado no programa *SPSS*.

Ainda de acordo com o *Teste de Correlação de Spearman*<sup>15</sup>, o mesmo grupo com menos exposição acadêmica e com treinamento demonstrou também correlação positiva, forte e significativa entre o que produziu dos heterotônicos nas frases-veículo e o que percebeu, diferentemente do pré-teste. Isso quer dizer que produzir adequadamente os heterotônicos inseridos nas frases-veículo esteve associado ao fato de perceber adequadamente esses mesmos heterotônicos em testes auditivos (e vice-versa). A Figura 5 mostra a tendência dessa associação positiva entre perceber e produzir as frases-veículo com heterotônicos. A forte relação entre as duas variáveis apontou que, dos dez participantes deste grupo, representados por um losango azul escuro no gráfico, apenas dois deles, que estão circulados, ficaram mais distantes da linha de tendência da correlação entre produção e percepção. A propensão em aumentar os participantes que tiveram forte correlação neste contexto pode estar associada ao fato de que, como já dito anteriormente, nas frases-veículo, os heterotônicos estão em destaque, o que pode facilitar a correlação entre percebê-los e produzi-los.



Fonte: Elaboração própria.

Figura 5 – Correlação entre produção das frases-veículo do grupo com treinamento e menos exposição acadêmica. *Scatterplot* rodado no programa *SPSS*.

Os outros três grupos, i.e., com menos exposição acadêmica e sem treinamento, e com mais exposição acadêmica e com e sem treinamento, não tiveram correlação significativa entre o que produziram e perceberam no pós-teste.

<sup>15</sup> Os valores do *Teste de Correlação de Spearman* são  $r=0,70$ ,  $p=0,024$ .

As correlações entre percepção e produção no pré-teste e também no pós-teste indicaram que os grupos com menos exposição acadêmica (com e sem treinamento) são os que fizeram em algum momento essa correlação, apoiaram-se nos dados de ambos os testes (percepção e produção) para tentar produzir e/ou perceber adequadamente mais heterotônicos. Isso se confirmou principalmente com o grupo com treinamento e menos experiência acadêmica que, após as sessões de treinamento perceptual, fez correlação entre percepção e os dois pós-testes de produção (das frases com contexto e das frases-veículo). Enquanto que, no pré-teste, essa associação ocorreu apenas na percepção e na produção das frases com contexto. O mesmo não aconteceu com os grupos com mais exposição acadêmica, refutando a nossa hipótese de que todos os grupos utilizariam de maneira semelhante a correlação entre percepção e produção ao longo dos testes. Parece que, justamente por terem sido expostos a mais horas em disciplinas de espanhol na universidade, e também pelo fato de já conhecerem alguns heterotônicos usados na pesquisa, esses dois grupos com mais exposição acadêmica não transformaram os acertos da percepção em acertos nos dois testes de produção, pois em nenhum momento houve correlação entre eles (tanto nos pré-testes como nos pós-testes). Este resultado dialoga também com as afirmações dos autores Flege e Bohn (2001) e Altmann (2006) e Kijak (2009), i.e., que nem sempre parece haver uma relação direta (de causa e efeito) entre o que se percebe e o que se produz. E isso, neste estudo, parece ser mais saliente nos grupos com mais exposição acadêmica.

Ainda no pós-teste, somente os resultados do grupo com treinamento e menos exposição acadêmica tiveram correlação tanto na produção das frases com contexto e na percepção como na produção das frases-veículo e na percepção. Passamos a demonstrar agora os resultados do teste de generalização.

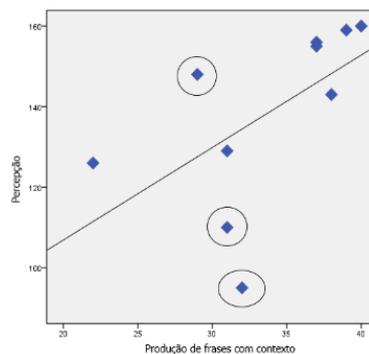
#### **4.3 Resultados da correlação entre os testes de generalização de produção e de percepção**

O *Teste de Correlação de Spearman* mostrou que nenhum dos quatro grupos dessa pesquisa fez associação entre o que produziu e percebeu (e vice-versa) nos testes de generalização. Relembramos que, nesses testes, apareceram novos locutores na percepção (não mais mexicanos, mas hondurenhos e cubanos) e também novos heterotônicos que ainda não tinham sido usados na pesquisa. Tal fato pode ter contribuído com esses resultados, pois, ao serem expostos a novos heterotônicos e a novos locutores, essa exposição a novos heterotônicos pode ter gerado uma ‘desorganização’ no sistema de aprendizagem, conforme preveem os Sistemas Complexos, o que fez com que os participantes desta pesquisa não utilizassem a percepção e a produção associadas como uma alternativa para tentar produzir e perceber adequadamente essas palavras heterotônicas.

Caspi e Lowie (2013) relatam que a recepção do vocabulário de uma L2 para a produção do mesmo está longe de ser linear. Isso é comprovado nesta etapa da pesquisa. Por outro lado, este estudo mostrou também que nem sempre a receptividade (na percepção) de itens do vocabulário é transformada para a produção, mesmo em um ambiente de grande variabilidade, ou seja, com novos heterotônicos ainda não vistos.

#### 4.4 Resultados da correlação entre os testes de retenção de produção e de percepção

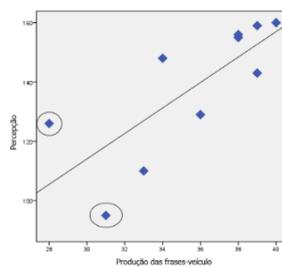
O grupo com treinamento e menos exposição acadêmica, segundo o *Teste de Correlação de Spearman*<sup>16</sup>, apresentou uma associação positiva, forte e significativa entre o que produziu dos heterotônicos nas frases com contexto e o que percebeu, nos testes de retenção. A forte relação entre as duas variáveis apontou que, dos dez participantes deste grupo, representados por um losango azul escuro no gráfico da Figura 6, sete seguiram próximos à linha de tendência da correlação e apenas três dos indivíduos (circulados) distanciaram-se mais dessa linha de tendência.



Fonte: Elaboração própria.

Figura 6 – Correlação entre produção das frases com contexto e percepção do grupo com treinamento e com menos exposição acadêmica. *Scatterplot* rodado no programa *SPSS*.

Segundo o *Teste de Correlação de Spearman*<sup>17</sup>, esse mesmo grupo possuiu ainda correlação positiva, forte e significativa entre o que produziu dos heterotônicos nas frases-veículo e o que percebeu. Isso quer dizer que produzir adequadamente os heterotônicos inseridos nas frases-veículo esteve associado ao fato de perceber adequadamente esses mesmos heterotônicos em testes auditivos (e vice-versa). A forte relação entre as duas variáveis apontou que, dos 10 participantes do grupo com treinamento e menos exposição acadêmica, representados por um losango azul escuro no gráfico da Figura 7, oito estão próximos à linha de tendência da correlação e apenas dois tenderam a afastar-se mais dessa linha.



Fonte: Elaboração própria.

Figura 7 – Correlação entre produção das frases-veículo e percepção do grupo com treinamento e com menos exposição acadêmica. *Scatterplot* rodado no programa *SPSS*.

<sup>16</sup> Os valores do *Teste de Correlação de Spearman* são  $r=0,70$ ,  $p=0,026$ .

<sup>17</sup> Os valores do *Teste de Correlação de Spearman* são  $r=0,84$ ,  $p=0,002$ .

Os outros três grupos desta pesquisa não tiveram correlação significativa entre o que produziram e perceberam nos testes de retenção. Assim, novamente os testes de correlação reforçaram que apenas o grupo com treinamento e menos exposição acadêmica parece ter sido beneficiado com as sessões de treinamento perceptual, além de ser o único grupo que manteve, ao longo dos testes (com exceção dos testes de generalização), a tendência de associar a percepção e a produção dos heterotônicos. Conforme previsto pelos Sistemas Dinâmicos (Larsen-Freeman, 2018), o sistema linguístico que está em desenvolvimento tende a apresentar um período de desestabilização quando algo novo (que normalmente não existe na L1) é apresentado ao aprendiz. Após um período de ‘reajuste’, esse novo sistema linguístico tende a se reorganizar. Para este grupo, parece que foi o que aconteceu. Nos pré-testes, nem sempre houve correlação entre percepção e produção, pois o sistema estava caótico diante de algo novo e diferente da L2 em relação à L1. Após os treinamentos perceptuais, a maioria dos participantes deste grupo passa a transformar a percepção adequada em produção adequada, auto-organizando o desenvolvimento do acento dos heterotônicos na L2.

Assim, os resultados desta pesquisa apontam que apenas aprendizes de níveis mais iniciais, que chamamos aqui de aprendizes com menos exposição acadêmica ao espanhol na universidade, utilizaram a percepção e a produção associadas na hora de desenvolver os heterotônicos e, ainda, que as sessões de treinamento perceptual ajudaram a fortalecer esta relação. Isso porque o grupo com menos exposição acadêmica e que não treinou não manteve, ao longo da pesquisa, a correlação entre produzir e perceber adequadamente as palavras heterotônicas, demonstrando que o sistema se manteve caótico por falta de entendimento de como funcionavam os heterotônicos, já que esse grupo não participou das sessões de treinamento. E os grupos com mais exposição acadêmica (o que treinou e o que não treinou perceptualmente) não mantiveram em nenhum momento essa correlação.

## 5 CONSIDERAÇÕES FINAIS

Esta pesquisa é resultado de um estudo com diversas etapas de treinamento de percepção dos heterotônicos do espanhol a aprendizes brasileiros. Como o PB e o espanhol são línguas de acento não previsível, a primeira parte do estudo (a do treinamento perceptual de Milan, 2019) comprovou que houve uma certa facilidade na percepção adequada dos heterotônicos. E que o treinamento perceptual foi positivo na produção dos mesmos. Esse artigo questionava se essa percepção transformaria a produção adequada dos heterotônicos por correlação. Assim, pelo perfil de nossos participantes, também nos questionamos se o tempo de exposição ao espanhol na universidade, por aprendizes brasileiros, interferiria no uso da correlação entre percepção e produção. O Quadro 1 mostra, resumidamente, os resultados desta investigação<sup>18</sup>.

---

<sup>18</sup> Os testes de generalização não foram reportados, porque não tiveram correlação.

Quadro 1 - Resultados das correlações entre todos os testes de percepção e de produção e com os quatro grupos de participantes.

	<b>Grupo com treinamento e menos exposição acadêmica</b>	<b>Grupo com treinamento e mais exposição acadêmica</b>	<b>Grupo sem treinamento e menos exposição acadêmica</b>	<b>Grupo sem treinamento e mais exposição acadêmica</b>
<b>Correlação entre produção e percepção nos pré-testes</b>	Entre percepção e produção das frases com contexto	sem	Entre percepção e frases-veículo	sem
<b>Correlação entre produção e percepção nos pós-testes</b>	Entre percepção e produção das frases com contexto e entre percepção e produção das frases-veículo	sem	sem	sem
<b>Correlação entre produção e percepção nos testes de retenção</b>	Entre percepção e os dois testes de produção	sem	sem	sem

Fonte: Elaboração própria.

É possível perceber que os *Testes de Correlação de Spearman* demonstraram que há uma tendência maior do uso de correlação entre percepção e produção no grupo com menos exposição acadêmica. Ou seja, há uma correlação entre perceber e produzir adequadamente os heterotônicos do espanhol para este grupo de aprendizes brasileiros, fato mantido em todas as etapas do treinamento, menos na generalização, sobretudo para o grupo que treinou. O grupo que não treinou e tinha menos exposição acadêmica só apresentou esta correlação no pré-teste (especificamente no de produção das frases-veículo e percepção). O grupo com treinamento e menos exposição acadêmica, apesar de não ter feito correlação nos pré-testes de percepção e de produção das frases-veículo, depois passa a fazê-la nos demais testes. O que nos leva a concluir que há uma tendência de os grupos com *menor exposição à língua alvo na academia* utilizarem mais a correlação entre percepção e produção do que grupos de níveis mais avançados.

Ou seja, para línguas de acento não previsível (como o espanhol e também o PB), o acento precisa ser especificado lexicalmente e necessita codificação na representação lexical das palavras (Altmann, 2006) e isso, segundo os resultados desta pesquisa, necessita ser feito sobretudo nos níveis iniciais de aprendizado. Além disso, conforme Caspi e Lowie (2013), os resultados dos níveis receptivo e produtivo do conhecimento interagem de forma complexa ao longo do tempo, às vezes competindo e às vezes apoiando-se mutuamente. Foi o que ocorreu com o grupo com menos exposição acadêmica e que recebeu treinamento, pois, nos pré-testes, ainda não realizava correlação entre percepção e produção das frases-veículo, mas depois passa a fazê-la. Vale ressaltar ainda que nem todos os aprendizes de cada grupo mantêm um mesmo padrão de correlação (o que já não era esperado por nós, justamente por nos filarmos à teoria dos Sistemas Complexos). Conforme as figuras demonstradas ao longo desta pesquisa, há sempre participantes que aparecem mais distantes da tendência de seu grupo.

**REFERÊNCIAS**

- Altmann H. The perception and production of second language stress: a cross-linguistic experimental study [tese]. Newark: Universidade de Delaware; 2006.
- Altmann H, Kabak B. Second language phonology. In: Kula N, Botma B, Nasukawa K, editores. *The continuum companion to phonology*. Nova Iorque: Continuum International Publishing Group; 2011. p. 298-319.
- Best T. A direct realist view of cross-language speech perception. In: Strange W, editor. *Speech perception and linguistic experience: theoretical and methodological issues in crosslanguage speech research*. Timonium, MD: York Press Inc; 1995. p. 171-206.
- Best T, Tyler M. Nonnative and second-language speech perception: commonalities and complementarities. In: Bohn S, Munro M, editores. *Second language speech learning: the role of language experience in speech perception and production*. Amsterdã: John Benjamins; 2007. p. 13-34.
- Caspi T, Lowie W. The dynamics of L2 vocabulary development: a case study of receptive and productive knowledge. *Revista Brasileira de Linguística Aplicada*. 2013;13(2):437-462.
- Colantoni L, Steele J, Escudero P. *Second language speech: theory and practice*. Cambridge: Cambridge University Press; 2015.
- Dupoux E, et al. A destressing 'deafness' in French? *Journal of memory and language*. 1997;36(3):406-421.
- Dupoux E, et al. Persistent stress 'deafness': the case of French learners of Spanish. *Cognition*. 2008;106(2):682-706.
- Dupoux E, Peperkamp S. Fossil markers of language development: phonological 'deafnesses' in adult speech processing. In: Laks B, Duran J, editores. *Phonetics, phonology and cognition*. Oxford: Oxford University Press; 2002. p. 168-190.
- Escudero P. *Linguistic perception and second language acquisition: explaining the attainment of optimal phonological categorization* [tese]. Utrecht: Universidade de Utrecht; 2005.
- Flege J. Second-language speech learning: Theory, findings and problems. In: Strange W, editor. *Speech perception and linguistic experience: theoretical and methodological issues in cross-language speech research*. Timonium, MD: York Press; 1995. p. 233-277.
- Flege J, Bohn O. The revised speech learning model (SLM-r). In: Wayland R, organizador. *Second language speech learning: theoretical and empirical progress*. Cambridge: Cambridge University Press; 2021. p. 3-83.
- Kijak A. *How stressful is L2 stress? A cross-linguistic study of L2 perception and production of metrical systems*. Utrecht: Universidade de Utrecht; 2009.
- Larsen-Freeman D. Task repetition or task interaction. In: Bygate M, editor. *Learning language through task repetition*. Amsterdã: John Benjamins Publishing Company; 2018. p. 311-330.
- Llisterri J. Relationships between speech production and speech perception in a second language. *Proceedings of the 13th International Congress of Phonetic Sciences*; 1995 August 13-19; Stockholm, Sweden; 1995. p. 92-99.
- Lucente L. *Aspectos dinâmicos da fala e da entoação no português brasileiro* [tese]. Campinas: Instituto de Estudos da Linguagem, Universidade Estadual de Campinas; 2012.
- Milan P. *Efeitos do treinamento perceptual na percepção e produção dos heterotônicos por aprendizes brasileiros de espanhol* [tese]. Curitiba: Universidade Federal do Paraná; 2019.
- Peperkamp S, Dupoux E. A typological study of stress 'deafness'. In: Gussenhoven C, Warner N, editores. *Laboratory Phonology 7*. Berlin: Mouton de Gruyter; 2002. p. 203-240.
- Rauber A, et al. TP: testes de percepção e treinamento perceptual com feedback imediato [programa de computador]. Versão 3.1. 2013.
- Rochet B. Speech perception and the phenomenon of foreign accent. *Journal of the Canadian Acoustical Association*. 1996;24(3):36.

## Desenvolvimento dos sons vocálicos do inglês como L2 por brasileiros em contexto de imersão

*Development of vowel sounds in English as an L2 by Brazilians in an immersion context*

Anderson Romário Souza Silva\*

*Universidade Federal Rural do Semi-Árido, Caraúbas, RN, Brasil*

Ronaldo Manguera Lima Júnior\*\*

*Universidade de Brasília, Brasília, DF, Brasil*

Lucas Heitor Ananias Oliveira\*\*\*

*Universidade do Estado do Rio Grande do Norte, Mossoró, RN, Brasil*

Dominyk Santos Dias Simões\*\*\*\*

*Universidade do Estado do Rio Grande do Norte, Mossoró, RN, Brasil*

**Resumo:** Esta pesquisa baseia-se na teoria de língua enquanto sistema dinâmico e complexo, apresentando uma discussão sobre atrito linguístico no português brasileiro (PB) e reorganização do sistema vocálico no inglês como segunda língua (IL2). Analisamos o efeito do contexto de imersão da segunda língua (L2) no desenvolvimento do sistema vocálico do PB e do IL2 através de uma pesquisa longitudinal com dados referentes a quatro coletas antes, durante e depois de um período de estadia de nove meses de 10 estudantes brasileiros nos Estados Unidos da América. Foram coletados valores de F1 e F2 e de duração vocálica de vogais nas duas línguas, a fim de analisar descritivamente o sistema vocálico dos falantes. Questionamos: qual o efeito da imersão na L2 no sistema vocálico de falantes brasileiros? Como hipótese, acreditamos que uma maior imersão na L2 resultará em um maior desenvolvimento do sistema vocálico do inglês, o qual pode gerar influências nas vogais do PB e estados de reorganização do sistema da L1. Os resultados indicam que houve reorganização das vogais do inglês-L2, mas não houve mudanças consideráveis nas vogais do PB.

**Palavras-chave:** Vogais. Imersão na L2. Atrito linguístico. Inglês como segunda língua. Português brasileiro.

**Abstract:** This study is based on the theory of language as a Dynamic and Complex System, discussing language attrition in Brazilian Portuguese (BP) and the reorganization of the vowel system in English as a second language. We analyzed the effect of the second language (L2) immersion context on the development of the vowel system in BP and English through a longitudinal study

---

\* Professor Adjunto, Departamento de Linguagens e Ciências Humanas, Universidade Federal Rural do Semi-Árido, Caraúbas, RN, Brasil; anderson@ufersa.edu.br

\*\* Professor Adjunto, Departamento de Linguística, Universidade de Brasília, Brasília, DF, Brasil; ronaldo.junior@unb.br

\*\*\* Mestrando em Ciências da Linguagem do Programa de Pós-Graduação em Ciências da Linguagem, Faculdade de Letras e Artes, Universidade do Estado do Rio Grande do Norte, Mossoró, RN, Brasil; lucasheitor3020@gmail.com

\*\*\*\* Mestranda em Ciências da Linguagem do Programa de Pós-Graduação em Ciências da Linguagem, Faculdade de Letras e Artes, Universidade do Estado do Rio Grande do Norte, Mossoró, RN, Brasil; dominykdias@alu.uern.br

with data from four collection points before, during, and after over approximately nine months of the stay of 10 Brazilian students in the United States. F1 and F2 values and vowel duration were collected in both languages to descriptively analyze the vowel system of the speakers. We question: what is the effect of L2 immersion on the vowel system of Brazilian speakers? As a hypothesis, we believe that greater immersion in the L2 will result in greater development of the English vowel system, which may influence BP vowels and cause reorganization of the L1 system. The results indicate that there was a reorganization of the L2-English vowels, but there were no significant changes in the BP vowels.

**Keywords:** Vowels. L2 Immersion. Language attrition. English as a second language. Brazilian Portuguese.

## 1 INTRODUÇÃO

O objetivo geral deste artigo é analisar o efeito do contexto de imersão na segunda língua (L2) no desenvolvimento do sistema vocálico do português brasileiro (PB) e do inglês. Para atingir esse objetivo, foi realizada uma pesquisa longitudinal que analisou dados referentes a quatro coletas ao longo de aproximadamente nove meses de estadia de 10 estudantes brasileiros residentes nos Estados Unidos da América (EUA). Foram coletados os valores de frequências formânticas (F1 e F2) e duração vocálica do PB e do inglês antes, durante e depois do período de imersão na L2, o qual foi quantificado através da aplicação de um questionário sociolinguístico (Anderson et al., 2018).

A pergunta-problema que objetivamos responder é: qual o efeito da imersão na L2 no sistema vocálico de falantes brasileiros? Como hipótese, acreditamos que uma maior imersão na L2 resultará em um maior desenvolvimento do sistema vocálico do inglês, o qual pode gerar influências nas vogais do PB e estados de reorganização do sistema (Schmid; Kopke, 2019). Assim, esperamos que o sistema vocálico da L2 seja beneficiado pelo período de imersão, mesmo após retorno ao país de origem. Entretanto, também prevemos que a variação na primeira língua (L1), se ocorrer, seja baixa e não permanente, caracterizando-se como deriva fonética, pois o sistema vocálico retornará ao seu estágio inicial após o fim do período de imersão (Chang, 2019).

Baseamos esta pesquisa na teoria dos Sistemas Dinâmicos e Complexos (SDC) (Larsen-Freeman, 1997; De Bot et al., 2007; Beckner et al., 2009), a qual propõe que a linguagem é um sistema complexo formado pela interação de diversos agentes que sofrem influências constantes e mútuas a partir de seu uso pelo falante, culminando na emergência de mudanças linguísticas tanto no nível da comunidade, quanto no nível individual. Com base nessa visão, as vogais no sistema vocálico da L1 atuam como atratores para a produção das vogais na L2, até que o indivíduo seja capaz de criar categorias distintas. Assim, entende-se que o estado inicial do sistema vocálico da L2 será o sistema vocálico da L1 e uma experiência de imersão na L2 pode gerar mudanças significativas, favorecendo a criação de novas categorias. Porém, a teoria propõe que, devido à linguagem ser não linear, também é possível que nenhuma mudança significativa ocorra, dependendo do contexto de desenvolvimento de cada indivíduo.

Essa visão de língua advoga influências mútuas entre diferentes sistemas vocálicos, permitindo que não apenas a L1 influencie o desenvolvimento da L2, mas também o contrário, caracterizando atrito linguístico (Schmid; Kopke, 2019). Atrito

linguístico é um termo generalista que se refere à influência de línguas adicionais na L1, as quais podem emergir devido a diversos fatores ou contextos (Schmid, 2011; Schmid; Kopke, 2019). Esse fenômeno é dinâmico, evidenciando que a L1 não está em estado fixo e pode sofrer constantes mudanças devido a fatores diferentes (Leeuw; Opitz; Lubinska, 2013). Porém, o contexto analisado nesta pesquisa refere-se a um tipo específico de atrito chamado de deriva fonética, pois acreditamos que as mudanças na L1 dos participantes são decorrentes de uma experiência recente de intercâmbio e não serão permanentes (Chang, 2019).

A literatura relata efeitos de deriva fonética até mesmo nos estágios iniciais de desenvolvimento da L2 e em estudantes recém-chegados no contexto da L2 (Chang, 2012, 2013). Chang (2013) também reporta que falantes com maiores níveis de proficiência na L2 mostram-se mais resistentes aos efeitos de deriva fonética comparados àqueles de menor proficiência, provavelmente devido ao segundo grupo estar sujeito a uma maior variação da L2 durante o período de aprendizagem. Por fim, os efeitos causados pela deriva fonética tendem a ser reversíveis após retorno a um contexto de imersão na L1, não causando mudanças duradouras (Chang, 2019).

Como exemplos de pesquisas voltadas para deriva fonética, semelhantes a este artigo, citamos os trabalhos de Lang e Davidson (2019), Kartushina e Martin (2019) e Turner (2022). Lang e Davidson (2019) reportam que não encontraram mudanças no sistema vocálico do inglês de falantes de inglês-L1/francês-L2 residentes nos EUA, mas identificaram um movimento de F1 nas vogais do inglês em direção às vogais do francês para falantes residentes em Paris. Kartushina e Martin (2019) analisaram a produção de vogais de falantes de espanhol-L1/inglês-L3 em um período de estudos em um contexto de imersão na L3. Os resultados mostraram um abaixamento considerável das vogais da L1, aproximando-se das vogais alvo da L3. Por fim, Turner (2022) analisou o sistema vocálico de falantes de inglês-L1/francês-L2 e constatou um alteamento das vogais da L1 à medida que os participantes tiveram maior variação na aprendizagem das vogais da L2, alinhando-se à noção de que falantes com menos proficiência na L2 estão mais sujeitos à deriva fonética (Chang, 2019).

Também é importante relatar que é possível haver deriva fonética na L1 em contexto de imersão na própria L1. Schereschewsky, Alves e Kupske (2017) reportaram essa influência bidirecional em brasileiros aprendizes de inglês como segunda língua (L2), os quais produziram o *voice onset time*<sup>1</sup> (VOT) de /k/ do PB com duração significativamente diferente de falantes monolíngues. Porém, acreditamos que o contexto de imersão na L2 favoreça, ainda mais, tais constatações.

O contexto de imersão na L2 é um fator importante para a emergência de atrito (Opitz, 2019), pois tende a inibir o uso da L1 (Linck; Kroll, 2019). Diferentes indivíduos podem estar sujeitos a diferentes níveis de atrito (ou atrito nenhum), mesmo que esses falantes tenham biografias semelhantes. Geralmente, mas não exclusivamente, indivíduos com maior nível de integração na comunidade falante da L2 e menos uso de sua L1 estão mais sujeitos ao fenômeno (Schmid, 2011; 2019). Assim, espera-se que os participantes que apresentem maior nível de imersão na L2 estejam mais propícios a desenvolverem o sistema vocálico do inglês, porém, também estejam sujeitos a um efeito maior de deriva fonética nas vogais do PB.

---

<sup>1</sup> Medida do tempo entre a liberação da plosiva e o início da vibração das cordas vocais do próximo segmento.

## 2 METODOLOGIA

Esta é uma pesquisa longitudinal que analisou o sistema vocálico de 10 participantes através de análise acústica de gravações coletadas ao longo de 11 meses.

### 2.1 Participantes

Os participantes desta pesquisa são 10 sujeitos que realizaram doutorado sanduíche em diferentes universidades dos EUA durante nove meses entre 2021 e 2022. A idade dos participantes variou de 27 a 37 anos, no início do intercâmbio. Todos eram estudantes de doutorado de diferentes áreas, programas de pós-graduação e instituições brasileiras, com níveis de proficiência variando entre B2 e C2, atestados através do TOEFL – IBT como requisito para participação no programa sanduíche. Os participantes são oriundos de diversos estados brasileiros (São Paulo, Rio Grande do Sul, Alagoas, Mato Grosso e Minas Gerais) e alguns reportaram, além do inglês como L2, proficiência em terceiras línguas (espanhol, francês ou italiano) e viagens prévias ao exterior. O grupo de participantes é composto por cinco falantes do gênero masculino e cinco do gênero feminino. Para preservar a anonimidade, nos referimos a todos os participantes utilizando o artigo masculino *o* e o número atribuído. Reconhecemos que os participantes não compõem um grupo homogêneo (considerando variantes da L1, línguas adicionais, local de residência e outras experiências de imersão), mas os dados aqui abordados formam um corpus importante e útil para a compreensão do tema estudado. Então, evitaremos fazer constatações gerais. Portanto, as constatações terão como foco os dados da pesquisa.

### 2.2 Experimento

O experimento utilizado consistiu na leitura de frases-veículo que continham duas palavras experimentais (ex: *a bola caiu no beco da vizinha / the cut looks bad today*). As vogais e palavras analisadas são apresentadas no Quadro 1. No total, foram analisadas sete vogais orais tônicas do PB e 12 do inglês, distribuídas em 19 frases. Cada *token* foi lido uma vez por coleta.

Quadro 1 - Vogais e palavras analisadas.

PORTUGUÊS						
/i/	/e/	/ɛ/	/a/	/ɔ/	/o/	/u/
pico	pelo	seta	taco	bola	toco	suco
dica	beco	beca	pata	toque	bolo	pulo
INGLÊS						
/i:/	/ɪ/	/ɛ/	/æ/	/ɔ/	/ɜ:/	

beat	kiss	guess	gas	pilot	shirt
sheep	ship	bed	bad	rocket	bird
/ʌ/	/a:/	/a/	/ɔ/	/ʊ/	/u:/
love	car	hot	talk	book	boot
cut	far	dog	shot	put	rude

Fonte: Elaborado pelos autores.

Cada participante produziu 38 vogais (14 do PB e 24 do IL2) por coleta, totalizando 152 vogais analisadas por indivíduo e 1520 no geral. A quantidade de palavras selecionadas para cada vogal foi uma limitação deste estudo, pois apenas duas produções de cada vogal por coleta não foram suficientes para mapear com precisão o espaço acústico ocupado por cada vogal no IL2, resultando em uma alta variação na distribuição das vogais. Acreditamos que aumentar o número de palavras por vogal ou a quantidade de repetições das palavras no experimento seria ideal, mas não foi possível implementar nesta pesquisa, pois os dados só foram analisados após todas as coletas terem sido concluídas.

### 2.3 Coleta de dados

Os dados foram coletados durante reuniões online individuais entre os pesquisadores e os informantes com o auxílio do software PRAAT versão 6.1.03 (Boersma; Weenink, 2021) para gravação (taxa de amostragem de 44kHz) nos próprios computadores dos informantes, com fones de ouvido em locais silenciosos. A adoção de tal método de coleta deu-se devido à impossibilidade de coleta presencial devido à pandemia e por causa de os participantes residirem em estados diferentes. Porém, a qualidade dos áudios foi satisfatória, como pode ser visto na Figura 1.

Foram realizadas quatro coletas ao longo de 11 meses, adaptando-se às datas diferentes de viagem de cada participante. Independentemente de quando o participante chegou nos EUA, a primeira coleta ocorreu na semana anterior à viagem, quando os participantes ainda estavam no Brasil imersos em um contexto falante de L1. A coleta dois ocorreu próxima do primeiro mês no contexto de imersão na L2. A coleta três aconteceu próxima do nono mês de imersão, o último do intercâmbio. Por fim, a coleta quatro ocorreu na primeira semana de regresso ao contexto de imersão na L1, após a conclusão do intercâmbio.

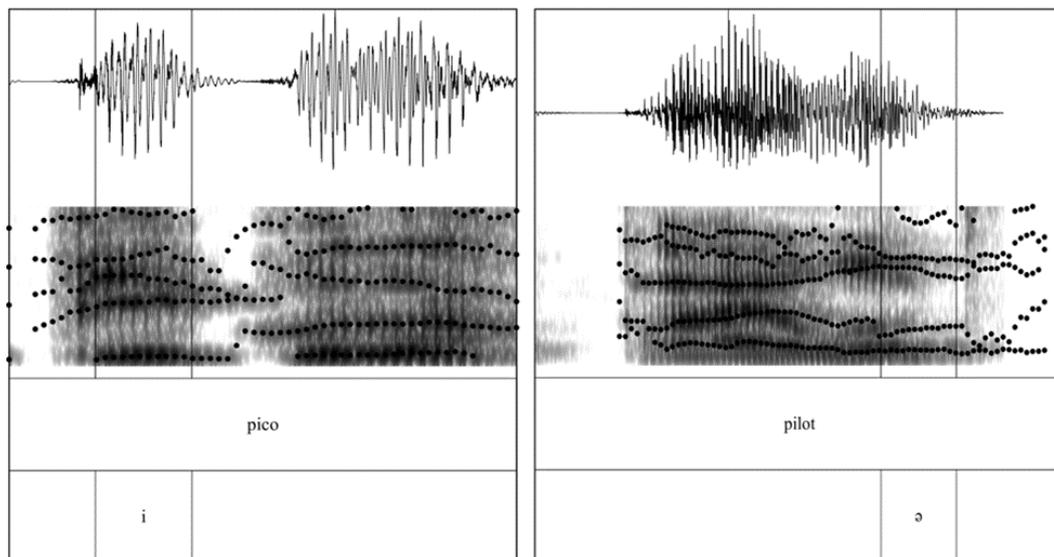
Inicialmente, foram planejadas coletas adicionais no terceiro e no sexto mês durante o intercâmbio (o que totalizaria seis coletas). Alguns dados dos participantes foram coletados nesses períodos, mas foram desconsiderados na análise por não estarem completos. Deste modo, há um espaço de tempo de cerca de sete meses no intercâmbio que não foi analisado. Esses dados seriam fundamentais para estudar o desenvolvimento do sistema vocálico desses participantes, porém as coletas não foram

concluídas devido a dificuldades pessoais do pesquisador. Acreditamos que as quatro coletas analisadas proporcionam uma análise válida do início e do fim da imersão na L2, bem como do retorno à imersão na L1, embora o meio desse percurso esteja ausente.

Em todas as coletas, foi realizado um questionário sociolinguístico adaptado de Anderson et al. (2018) para averiguar o nível de imersão na L2 dos participantes. O objetivo do questionário foi atribuir uma nota linear (de 0 a 195) para mensurar o quão imersos estavam os participantes na L2, considerando a pontuação obtida em suas respostas. O valor linear possibilita averiguar o nível de imersão na L2, evitando uma classificação categórica como “imerso/não imerso”. Essas perguntas foram relacionadas ao uso de ambas as línguas, em diferentes contextos. As gravações das duas línguas aconteceram em sequência. Uma conversa de cerca de 15 minutos em português, sobre temas diversos, foi realizada antes da coleta da L1. Por sua vez, o questionário foi aplicado em formato de entrevista em inglês para que os participantes entrassem em modo bilíngue, antes da coleta da L2.

## 2.4 Análises

A primeira etapa da análise dos dados foi uma análise acústica através do PRAAT versão 6.3.10 (Boersma; Weenink, 2023) a fim de obter os valores de F1 e F2 de cada vogal, os quais foram convertidos usando o método de Lobanov para calcular uma distância euclidiana entre as vogais, evitando, principalmente, a grande variação comum nos valores de F2 (Lima Jr., 2016). Também foram coletadas as durações das palavras e das vogais para obtenção dos valores da duração da vogal relativa à palavra. A Figura 1 apresenta dois exemplos de produções do Participante 3 das palavras *pico* e *pilot*, respectivamente, com o contorno formântico e as durações destacadas.



Fonte: Elaborado pelos autores.

Figura 1 - Produções de *pico* e *pilot* exemplificando, a análise acústica.

Após a análise acústica, os sistemas vocálicos dos participantes foram plotados, com o uso do pacote PhonR (McCloy, 2022) no R (R Core Team, 2023), permitindo uma análise descritivo-qualitativa do desenvolvimento dos sistemas e do movimento ou criação de novas categorias vocálicas. Acreditamos que a distribuição das vogais do

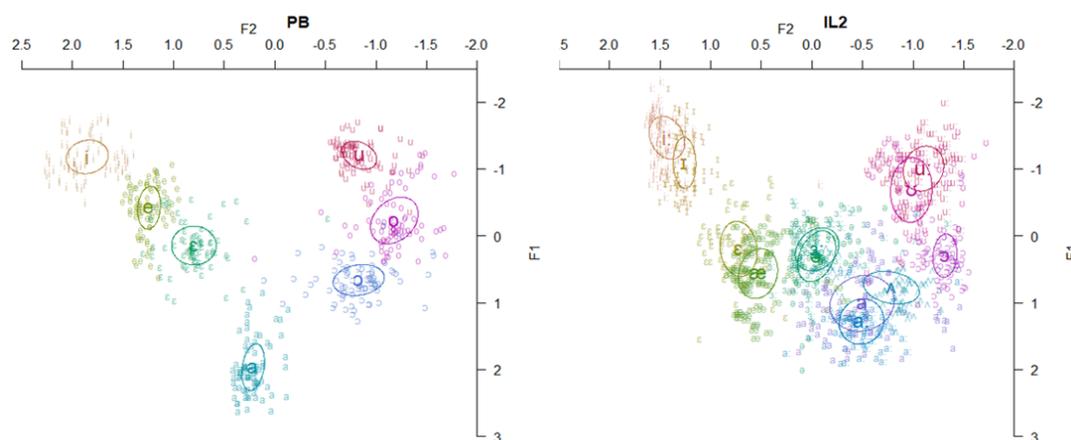
inglês aproxime-se do alvo à medida que a imersão na L2 aumente, o que pode ocasionar períodos de variação nas vogais da L1 durante o percurso de desenvolvimento da L2. Também esperamos que o retorno à imersão na L1 desfavoreça a distinção das vogais do IL2, ocasionando um retorno ao estado inicial do sistema vocálico, podendo haver benefícios ou não, dependendo da formação de novas categorias pelos participantes.

### 3 RESULTADOS

Os resultados desta pesquisa são apresentados em duas etapas. A primeira discute os achados gerais relacionados às mudanças no sistema vocálico dos participantes como um grupo, enquanto a segunda foca no desenvolvimento individual de cada participante. Ambas as etapas apresentam uma análise descritivo-observacional do espaço acústico médio ocupado por cada vogal, buscando delimitar as categorias vocálicas dos participantes para fins de comparação.

#### 3.1 Análise geral

A Figura 2 apresenta a distribuição vocálica do PB e do IL2 de todos os participantes, considerando as quatro coletas. Em cada língua, uma cor diferente foi atribuída para cada vogal. Os pontos representados por símbolos fonéticos menores representam cada ocorrência das respectivas vogais, enquanto os pontos ocupados pelos símbolos fonéticos maiores representam os valores médios de F1 e F2 de cada vogal. As elipses correspondem ao espaço acústico médio ocupado por cada vogal, considerando o intervalo de confiança de 95%.



Fonte: Elaborado pelos autores.

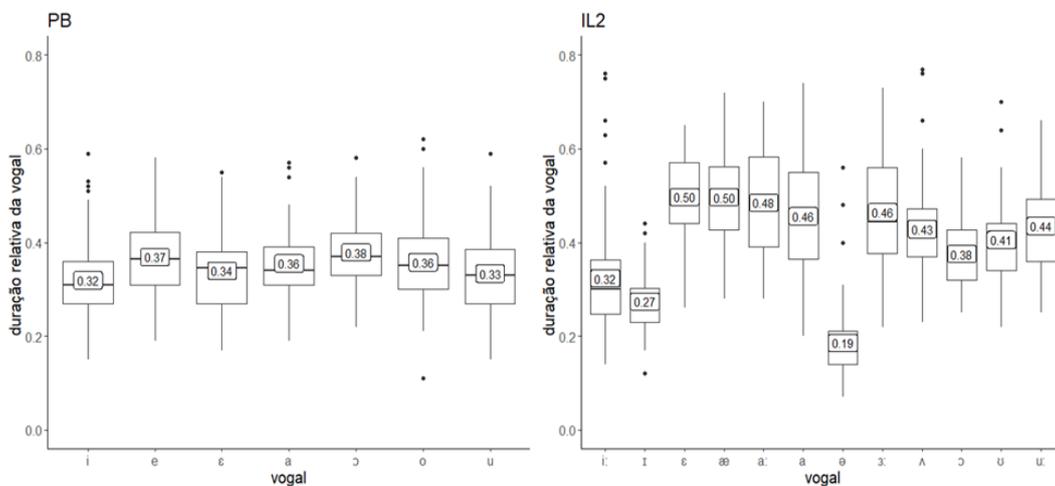
Figura 2 - Distribuição das vogais do PB (à esquerda) e do IL2 (à direita).

Em relação ao sistema vocálico do PB, é possível perceber que as vogais ocupam posições esperadas, sem nenhuma sobreposição nas elipses. Nota-se, também, uma alta variação nas produções de cada *token*, gerando áreas de sobreposições entre vogais vizinhas (como /i/ e /e/, por exemplo). Acreditamos que essa variação seja decorrente da natureza do corpus analisado, o qual foi composto por 10 falantes de diferentes variedades do PB em diferentes estágios de imersão no IL2 e produzindo palavras em sentenças diferentes, o que impossibilitou uma maior homogeneidade na distribuição das vogais devido ao desenvolvimento das duas línguas.

Essa variação é mais evidente no sistema vocálico do IL2, onde as vogais ocupam um espaço acústico maior do que as vogais do PB. Essa variação, nesse caso, é esperada, pois se trata da L2 dos participantes, mostrando que as categorias vocálicas não são tão homogêneas como na L1. Cada vogal aparenta ocupar o espaço acústico próximo do esperado, porém, observam-se várias sobreposições entre vogais próximas (como /ε/ e /æ/), mostrando que, no geral, o grupo de participantes apresenta problemas em distinguir vogais vizinhas da L2.

A Figura 3 apresenta a duração relativa das vogais em relação à palavra do PB e do IL2 de todos os participantes nas quatro coletas. No PB, os valores de duração em milissegundos das vogais foram: 94 /i/, 100 /e/, 130 /ε/, 105 /a/, 112 /ɔ/, 118 /o/, 112 /u/. A literatura relata que vogais altas tendem a ter menor duração do que vogais baixas devido à menor abertura do trato vocálico durante sua produção (Cristófaro-Silva, 2019). Convertendo os valores para duração relativa, as vogais altas /i, u/ e média-altas /e, o/ comportaram-se como esperado, porém, as vogais média-baixas /ε, ɔ/ e baixa /a/ apresentaram duração relativa inferior às primeiras. Novamente, a natureza do corpus pode justificar essa variação.

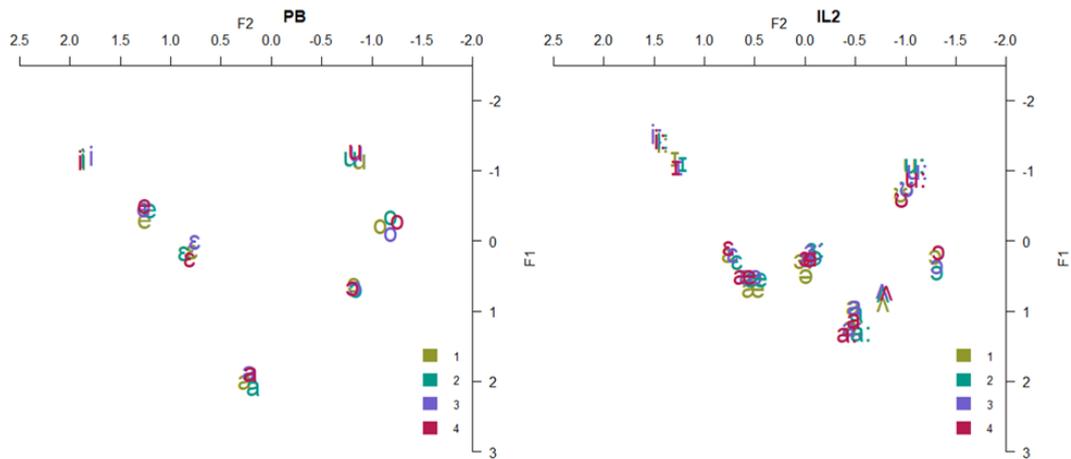
Nos dados do IL2, nota-se que o grupo tem dificuldades em distinguir algumas vogais em pares problemáticos para aprendizes brasileiros (como /ε/ e /æ/), mas não nos pares /i:/ e /ɪ/ ou /u:/ e /ʊ/, sendo a vogal longa produzida com maior duração relativa média. Também se percebe que o grupo foi capaz de produzir, satisfatoriamente, a vogal /ɔ/ com duração baixa, enquanto sua contraparte /ɜ:/ foi produzida com maior duração.



Fonte: Elaborado pelos autores.

Figura 3 - Duração relativa das vogais do PB (à esquerda) e do IL2 (à direita).

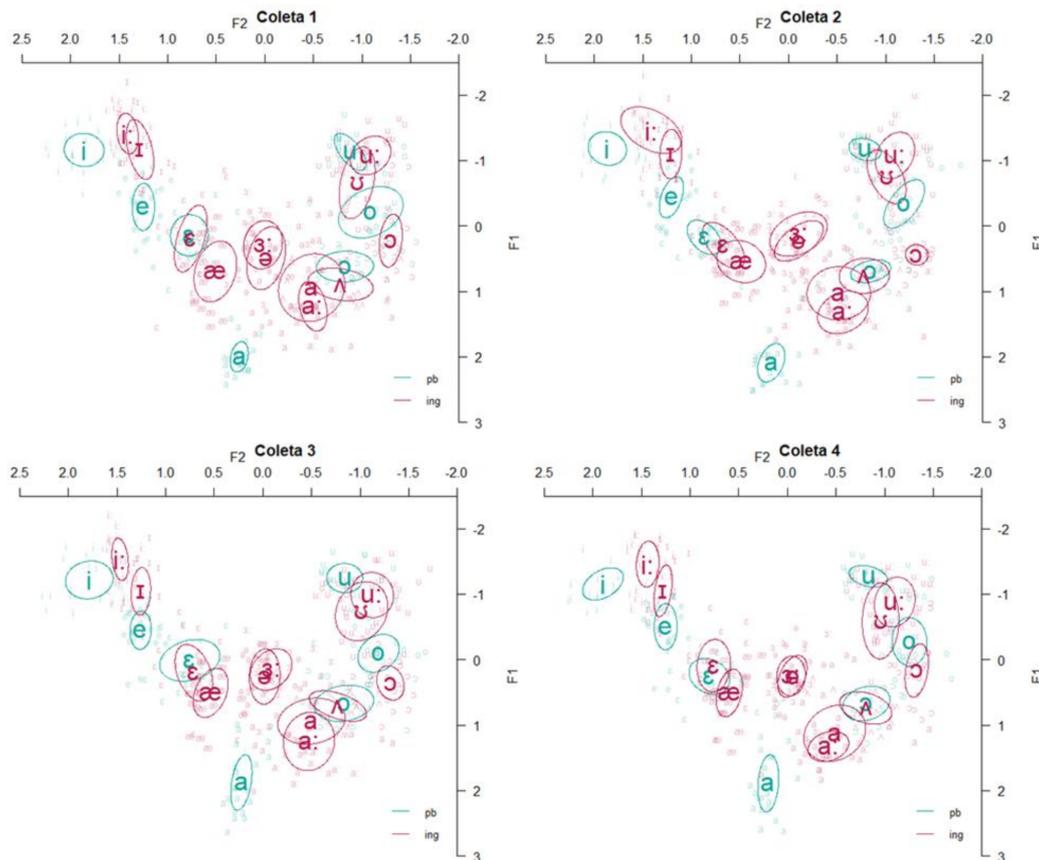
Os dados apresentados nas Figuras 2 e 3 oferecem um ponto de vista geral do sistema vocálico do grupo como um todo, mas não de seu desenvolvimento ao longo da imersão. Para este fim, a Figura 4 apresenta o desenvolvimento do grupo ao longo das quatro coletas, diferenciadas por cor.



Fonte: Elaborado pelos autores.

Figura 4 - Valores médios das vogais do PB (à esquerda) e do IL2 (à direita) por coleta.

Em ambas as línguas, não houve grandes movimentos no sistema vocálico ao longo das quatro coletas, considerando os valores médios de cada categoria vocálica. Todas as vogais mostraram-se homogêneas e a imersão no IL2 não resultou em movimentos inesperados, mesmo no ápice da imersão durante a coleta 3. A Figura 5 apresenta o sistema vocálico do grupo (nas duas línguas, diferenciadas por cores) em cada ponto de coleta. Essas plotagens permitem averiguar se houve algum movimento da L1 ou da L2 em direção à outra língua.



Fonte: Elaborado pelos autores.

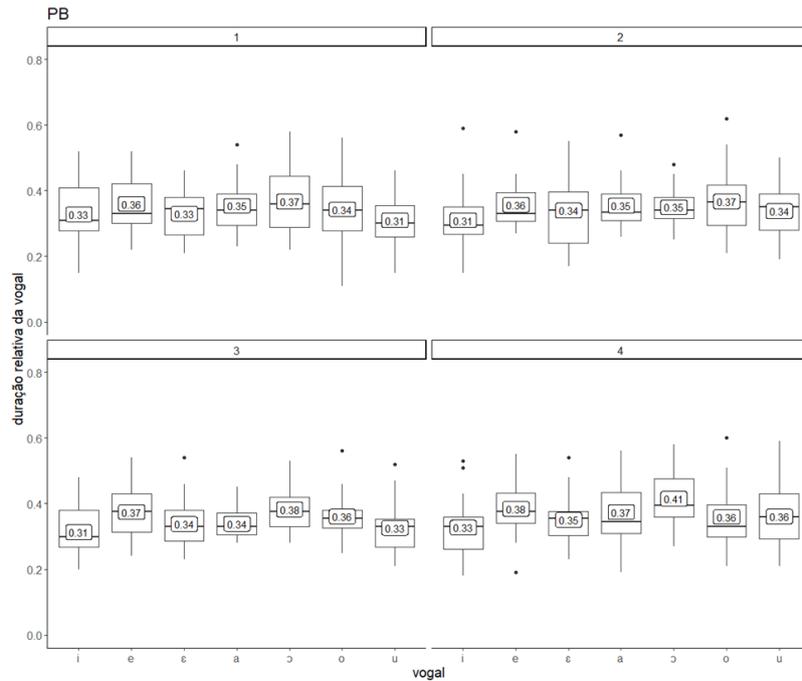
Figura 5 - Distribuição das vogais do PB (à esquerda) e do IL2 (à direita) por coleta.

Em todas as coletas, percebe-se que as vogais do PB estão bem delimitadas e não há sobreposições das elipses, como esperado para a L1. Também não há grandes movimentos das vogais da L1 ao longo das coletas, apenas nota-se que algumas vogais (como /a/) obtiveram elipses maiores nas últimas coletas, indicando maior variação no auge da imersão (coleta 3) e após o retorno (coleta 4).

Assim como no gráfico geral, o IL2 apresenta sobreposição de vogais vizinhas em todos os pontos de coleta e houve movimentos individuais nas vogais, mas não no sistema vocálico como um todo. Algumas vogais na L2 dividem o mesmo espaço acústico de vogais da L1 (como /ɛ/ nas duas línguas), mas percebe-se que o grupo distingue as vogais /i:/ e /u:/ do IL2, sobrepondo a influências das respectivas vogais atradoras próximas do PB. Também se nota que, ao longo de todas as coletas, o grupo produziu a vogal /ʌ/ do IL2 próxima da vogal /ɔ/ da L1, assim como as vogais /a, a:/ do IL2 foram produzidas aproximando-se da área central do espaço acústico, mais alta do que a vogal /a/ do PB. Antes da imersão, o grupo foi capaz de distinguir /ɛ/ e /æ/ satisfatoriamente, porém, há sobreposição durante a imersão, sendo reduzida após o retorno à imersão na L1 na coleta 4. Esse achado é explicado pela língua enquanto SDC, a qual propõe que há períodos de alta variabilidade no sistema (como durante a imersão), nos quais o falante produz maior esforço para favorecer uma característica específica (como aprender a distinção entre outros pares de vogais) e acaba desfavorecendo (momentaneamente) características já desenvolvidas (Beckner et al., 2009). Os pares de vogais /ɔ/-/ɜ:/ e /a/-/a:/ perdem a pouca distinção que tinham

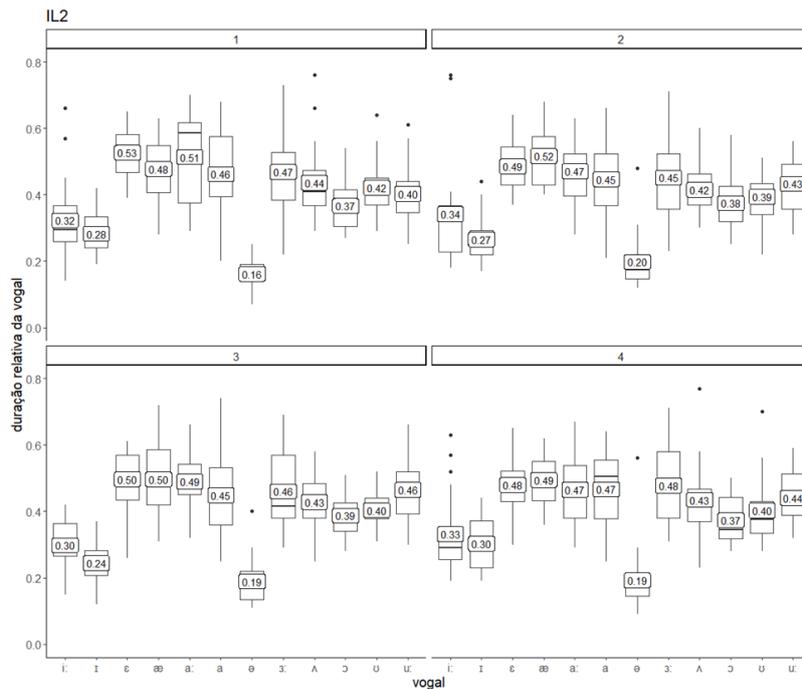
na coleta 4. No ápice da imersão (coleta 3), o grupo foi capaz de distinguir o par /i:/ e /I/ e manteve essa distinção após a imersão.

Considerando a análise descritiva, não há grande variação na duração relativa das vogais do PB antes ou após o período de imersão (Figura 6). Porém, há aumento na duração relativa de algumas vogais no IL2 nas coletas em imersão (Figura 7).



Fonte: Elaborado pelos autores.

Figura 6 - Duração relativa das vogais do PB por coleta.



Fonte: Elaborado pelos autores.

Figura 7 - Duração relativa das vogais do IL2 por coleta.

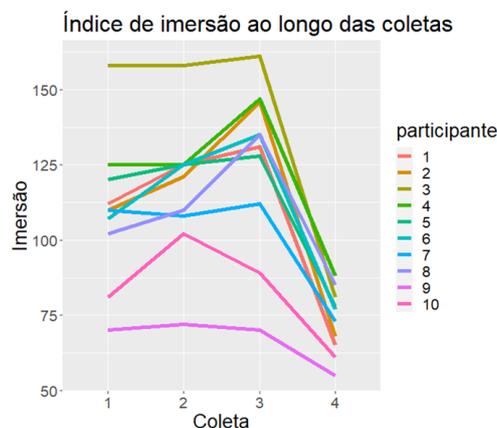
A vogal /ɪ/ é produzida com maior duração no retorno à imersão na L1, não diferenciando-a mais da vogal /i:/. Também há um aumento na duração da vogal /æ/ na coleta 2 que se manteve até a coleta 4. A distinção entre /u:/ e /ʊ/ foi produzida a partir da coleta 2. Um aumento considerável é visto na duração relativa de /ə/ a partir do início da imersão, porém, mantendo-se inferior à sua contraparte longa. A vogal /a:/ também sofre uma redução em sua duração relativa na coleta 4, perdendo a distinção com /a/.

A descrição dos dados apresentados nesta seção demonstra que não há vestígios marcantes de atrito linguístico nas vogais do PB antes, durante ou após a imersão na L2. Também revelam que as vogais da L2 apresentaram maior movimento no espaço acústico, às vezes em direção à formação de novas categorias ou à perda delas. Entretanto, uma análise descritiva geral não permite averiguar se houve mudanças significativas no sistema vocálico dos participantes no nível intraindividual. Para atingir esse objetivo, a próxima seção apresenta uma análise do desenvolvimento do sistema vocálico de cada participante.

### 3.2 Análise intraindividual

Esta seção apresenta uma análise descritiva-qualitativa do desenvolvimento do sistema vocálico dos participantes ao longo dos quatro pontos de imersão (uma semana antes da imersão, primeiro mês durante a imersão, nono mês durante a imersão, uma semana após a imersão), relacionando com as informações coletadas no questionário sociolinguístico. Os gráficos de distribuição vocálica apresentam o espaço acústico médio ocupado por cada vogal analisada, pois não foi possível gerar elipses individuais com a quantidade de pontos analisados por coleta. Esta é uma análise descritiva-observacional que considera a posição ocupada pelas vogais no espaço acústico para comparar suas movimentações ao longo das coletas.

A Figura 8 apresenta os valores de imersão de cada participante obtidos através do questionário sociolinguístico, representando o quão imerso o participante está na L2 naquele ponto. Em situações ideais e generalizadas, espera-se que o índice de imersão aumente à medida que o participante passe mais tempo imerso no IL2, pois o indivíduo estaria utilizando mais a L2 do que a L1, sendo reduzido esse índice após o retorno ao Brasil. Porém, os dados indicam que há variação intraindividual no nível de imersão, alinhando-se à proposta de língua enquanto SDC (Beckner et al., 2009).



Fonte: Elaborado pelos autores.

Figura 8 - Índices de imersão na L2 dos participantes.

Os índices de imersão da maioria dos participantes foram ao encontro do esperado. O valor na primeira coleta foi consideravelmente alto, devido à preparação para o intercâmbio. Os participantes revelaram diferentes formas de estudo do inglês, como aulas particulares e cursos preparatórios, além de engajamento com a L2 diariamente através de mídias digitais. O índice aumentou nas coletas seguintes, atingindo seu ápice no nono mês do intercâmbio. Um alto índice de imersão na L2 é um dos fatores que favorecem mudanças no sistema da L1, pois pode ser reflexo de um baixo uso da L1 (Linck; Kroll, 2019; Opitz, 2019).

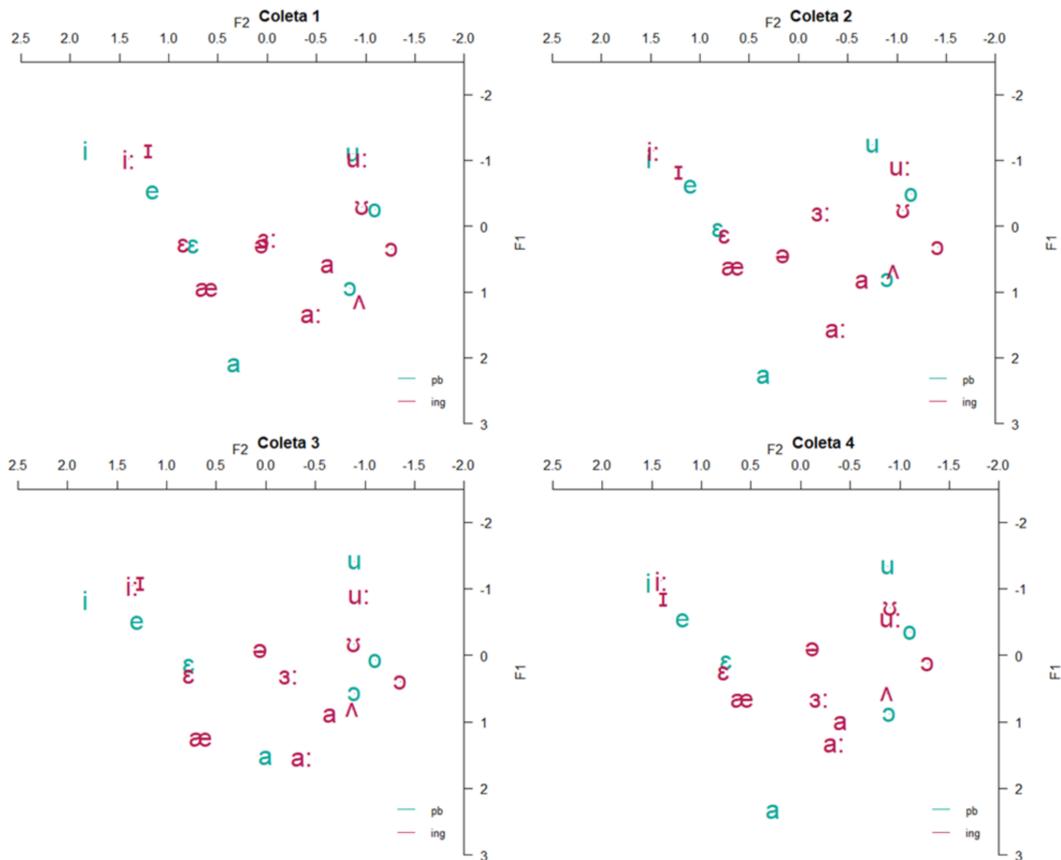
O contexto de cada participante foi diferente, contribuindo para índices distintos. Por exemplo, o Participante 7 obteve índices de imersão relativamente altos, que não aumentaram ao longo das coletas, pois residiu em uma comunidade de brasileiros na Flórida que desfavoreceu uma maior imersão na L2. Os Participantes 9 e 10, por sua vez, relataram um índice de imersão menor antes do intercâmbio (embora maior que o valor após imersão, devido aos seus estudos) que aumentou na coleta 2, mas regrediu na coleta 3. Esses dois participantes mudaram-se para o local do intercâmbio com seus cônjuges, estabelecendo um núcleo familiar que inibiu o uso da L2 diariamente. Além disso, houve uma queda brusca na imersão do participante 10 na terceira coleta, justificada devido à mudança para a modalidade remota de seu trabalho na universidade.

Todos os participantes relataram seus menores índices de imersão na L2 após o retorno ao Brasil, como esperado. No momento da coleta 4, alguns relataram cansaço em utilizar o inglês constantemente e, a maioria relatou que utilizava a L2 apenas quando necessário, para consumo de mídias ou comunicação escrita.

Passemos agora à análise intraindividual.

O Participante 1 é natural de Minas Gerais, tinha 29 anos durante a pesquisa e residiu no estado de Nova York. Relatou uso da L2 diariamente, obtendo um índice de imersão que aumentou consideravelmente ao longo do intercâmbio. No momento da segunda coleta, o participante relatou poucas oportunidades de contato social, pois morava sozinho. Mas, na terceira coleta, seu índice de imersão aumentou. O participante relatou que, naquele momento, teve mais chances de conhecer novas pessoas, pois estava morando com amigos.

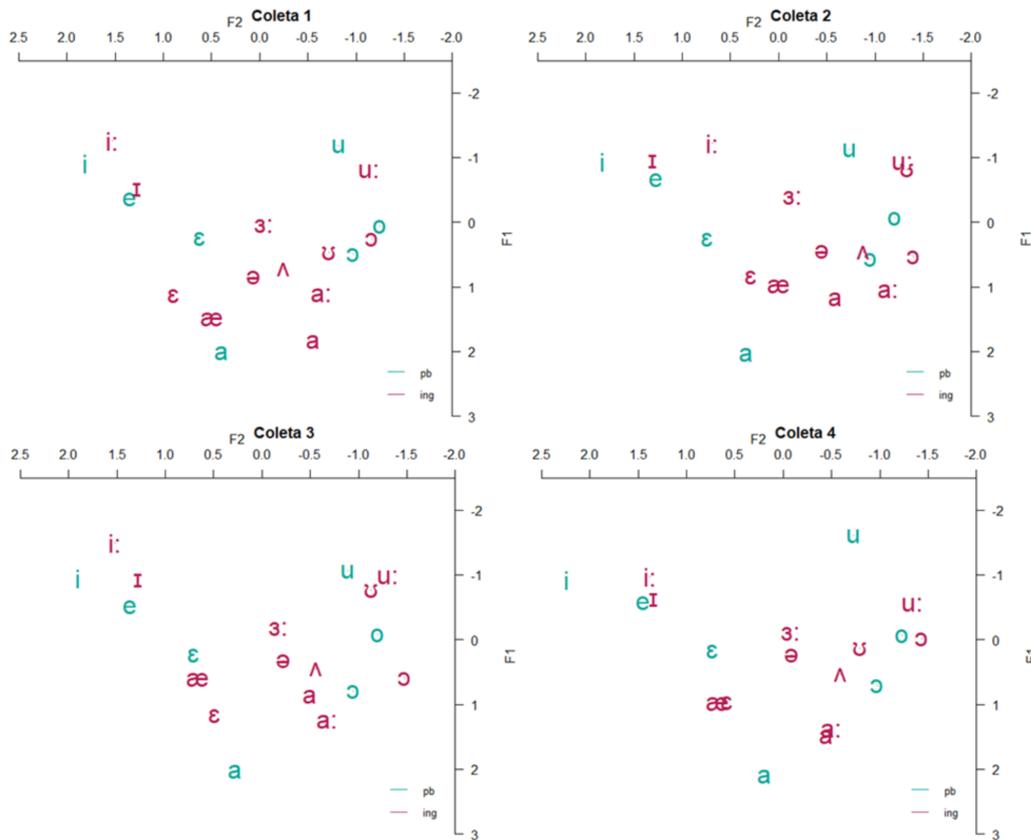
O sistema vocálico do PB manteve-se consistente durante as quatro coletas, embora seja possível perceber um movimento de /i/ em direção às vogais semelhantes do inglês na coleta 2 (o qual foi desfeito no nono mês) e na coleta 4. A vogal /a/ do PB também teve um abaixamento no primeiro mês da imersão, mas aproximou-se das vogais /a:/ e /a/ da L2 no nono mês, antes de retornar ao estado inicial após a imersão. Na coleta 4, percebe-se que o participante foi capaz de manter o contraste adquirido durante a imersão para as vogais centrais curtas (/ɔ/ e /a/), diferenciando-as de suas contrapartes longas. Porém, perdeu a distinção do par do IL2 devido à atração da vogal /i/ do PB, embora já tivesse demonstrado dominar essas categorias anteriormente, na primeira coleta. Assim, percebe-se uma reorganização dinâmica do sistema vocálico, favorecendo mudanças em certas categorias, enquanto outras perdem seu contraste (Beckner et al., 2009).



Fonte: Elaborado pelos autores.

Figura 9 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 1.

O Participante 2 é natural de São Paulo, tinha 34 anos durante a pesquisa e residiu em diversos estados durante a realização do intercâmbio, visitando instituições públicas para sua pesquisa. Relatou uso diário da L2 para realizar a sua pesquisa e alto engajamento social. Também obteve índice de imersão que aumentou ao longo da imersão, embora tenha mantido contato diário com familiares por chamadas na L1. No momento da coleta 2, o participante residia na Califórnia e teve contato com vários brasileiros, inclusive com o Participante 8. O aumento expressivo no índice de imersão na coleta 3 ocorreu enquanto residia em Nova York, tendo pouco contato com brasileiros.

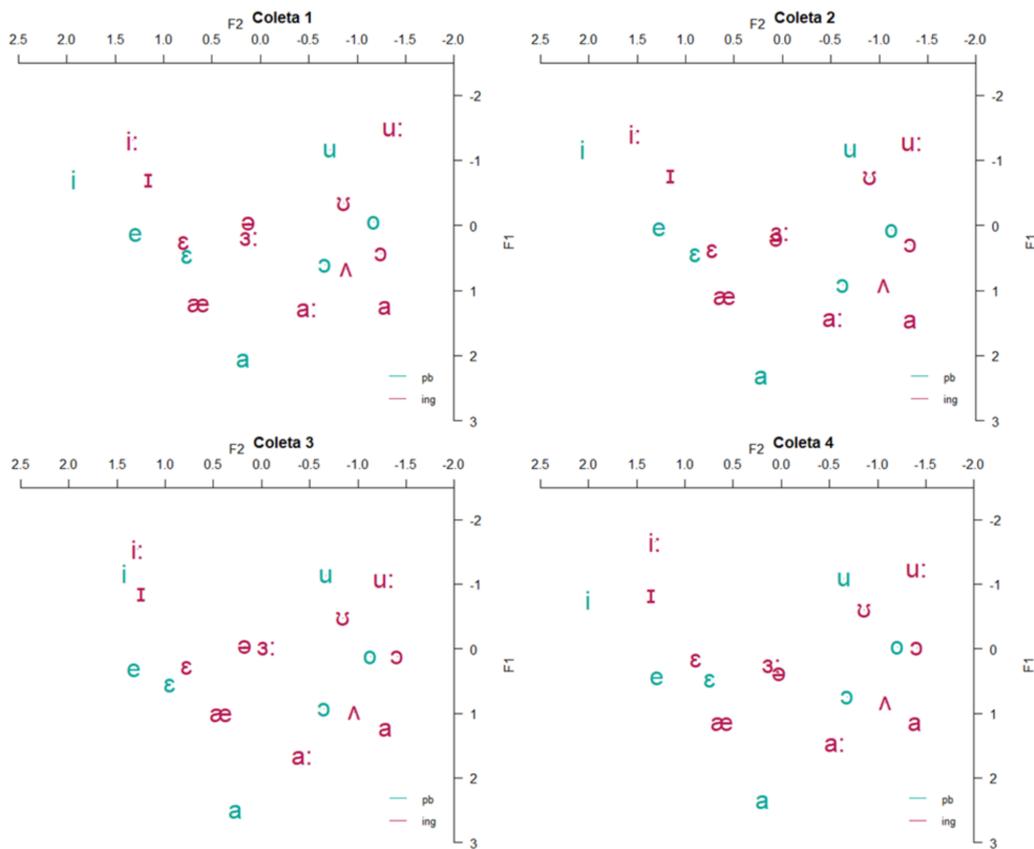


Fonte: Elaborado pelos autores.

Figura 10 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 2.

É possível perceber um recuo na produção da vogal /i/ da L1 na coleta 2, indo em direção à vogal semelhante /i:/ da L2, caracterizando uma mesma categoria para ambas. Esse movimento não ocorreu no ápice da imersão, embora tenha retornado na quarta coleta. Esse participante aparenta produzir as três vogais dessa região (/i/ do PB, /i:/ e /ɪ/ do IL2) com as mesmas categorias acústicas. As vogais /ɛ/ e /a/ da L1 permaneceram estáveis durante todas as coletas. Porém, as vogais /ɛ/ e /æ/ da L2, que aparentavam ocupar espaços distintos na primeira coleta, perderam seu contraste na coleta 4. A coleta 3 também aproximou as vogais /o/ e /ɔ/ na L1. Esse participante teve dificuldades em produzir algumas vogais do IL2 nos espaços acústicos esperados, tendo influência do sistema da L1, e, de maneira geral, houve uma perda no contraste das vogais no sistema da L2, após o retorno ao Brasil.

O Participante 3 é natural do Rio Grande do Sul, tinha 32 anos no início da pesquisa e obteve os mais altos índices de imersão do grupo. Este participante residiu em Illinois e reportou o inglês como sua língua principal durante o intercâmbio. Dividiu moradia com falantes de inglês como L1 e um brasileiro com quem se comunicava exclusivamente em IL2, trabalhou diariamente em um laboratório na universidade, onde se comunicava em inglês, e reservou o português apenas para o contato com familiares semanalmente e para uma visita do Participante 2.

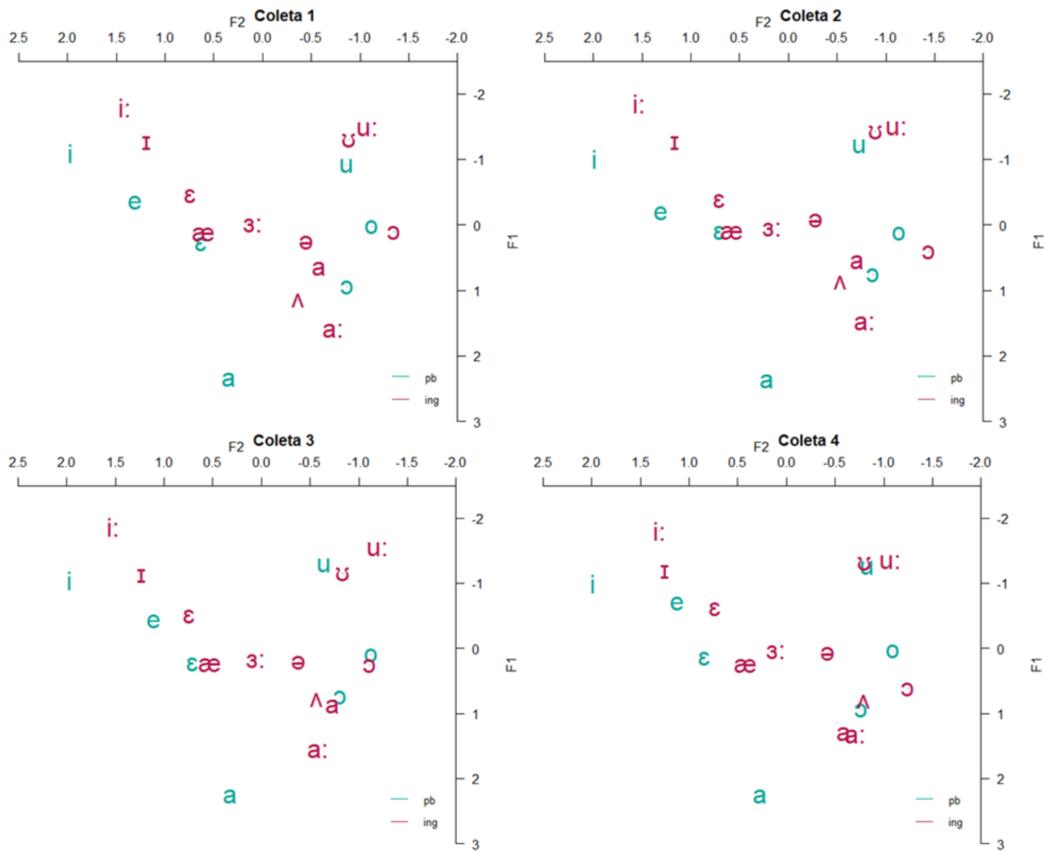


Fonte: Elaborado pelos autores.

Figura 11 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 3.

Na coleta 3, houve um movimento da vogal /i/ do PB em direção às categorias das vogais /i:/ e /ɪ/ do IL2, o qual foi desfeito após a imersão. As outras vogais da L1 permaneceram estáveis durante as quatro coletas. No IL2, o participante demonstrou dominar bem as categorias vocálicas do inglês, distinguindo satisfatoriamente as vogais no espaço acústico. Apenas o par /ə/ e /ɜ:/ teve sobreposição, sendo que a vogal longa é tipicamente produzida mais baixa. Não foram observadas mudanças drásticas na L1, embora esse participante seria, em tese, o candidato ideal para atrito linguístico, pois obteve o maior índice de imersão em todas as coletas.

O Participante 4 é natural do Rio Grande do Sul, tinha 29 anos na primeira coleta e residiu em New Jersey. Seu índice de imersão permaneceu o mesmo nas duas primeiras coletas e aumentou consideravelmente na terceira coleta. O participante justificou seu maior índice de imersão na terceira coleta devido à formação de um grupo de amigos e ao início de um relacionamento. A comunicação resultante desse relacionamento com um falante nativo de inglês fez com que o Participante 4 obtivesse o maior índice de imersão, após retorno ao Brasil, do grupo.

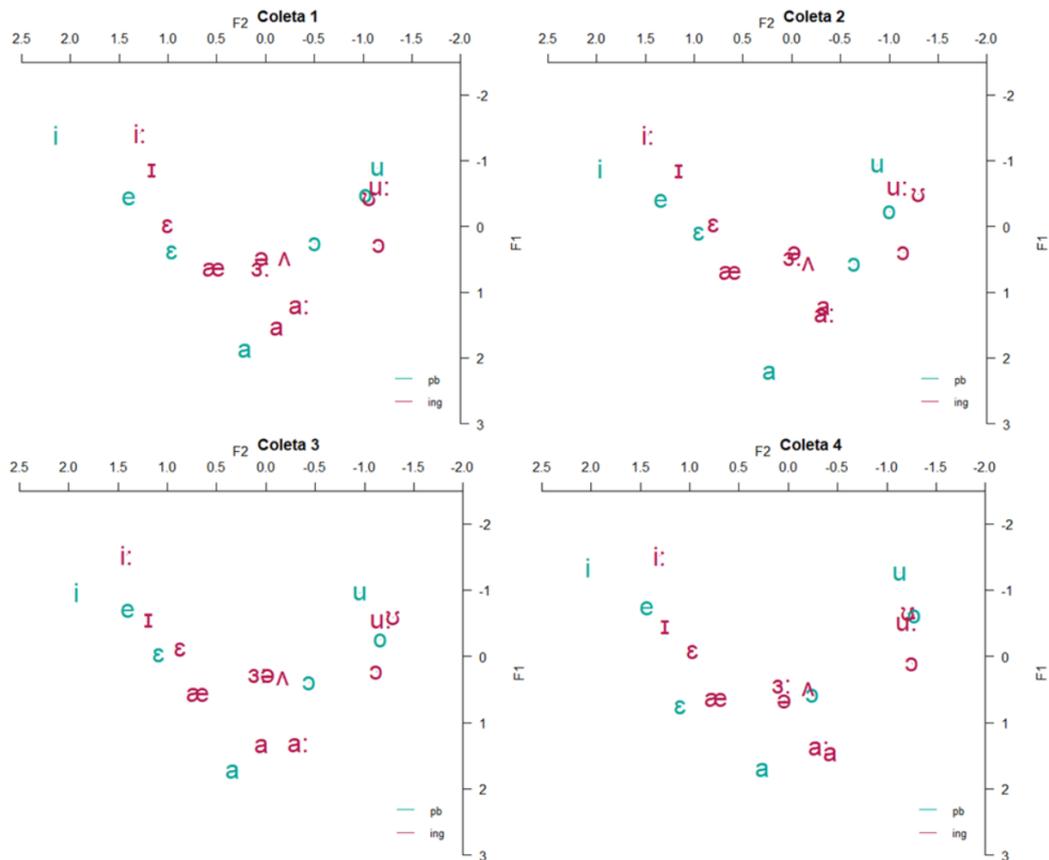


Fonte: Elaborado pelos autores.

Figura 12 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 4.

A L1 do Participante 4 permaneceu estável ao longo das quatro coletas, mas houve uma variação na produção da vogal /u/ do PB, parecendo formar uma mesma categoria para as vogais /u:/ e /ʊ/ da L2, após retorno ao Brasil. A terceira coleta foi o momento em que as vogais do IL2 foram melhor distribuídas no sistema, atingindo a distinção de /u:/ e /ʊ/, até então inexistente. Também houve perda da distinção de /a/ e /a:/ na L2 após o retorno ao Brasil, além da perda de distinção entre as vogais /u:/ e /ʊ/.

O Participante 5 é natural de Minas Gerais, tinha 28 anos na primeira coleta e residiu em Maryland. Residiu com um casal de falantes de inglês como L1 e relatou uso do inglês diariamente na universidade, mas pouco engajamento social (mesmo no nono mês de intercâmbio). Obteve um índice de imersão semelhante nas três primeiras coletas.

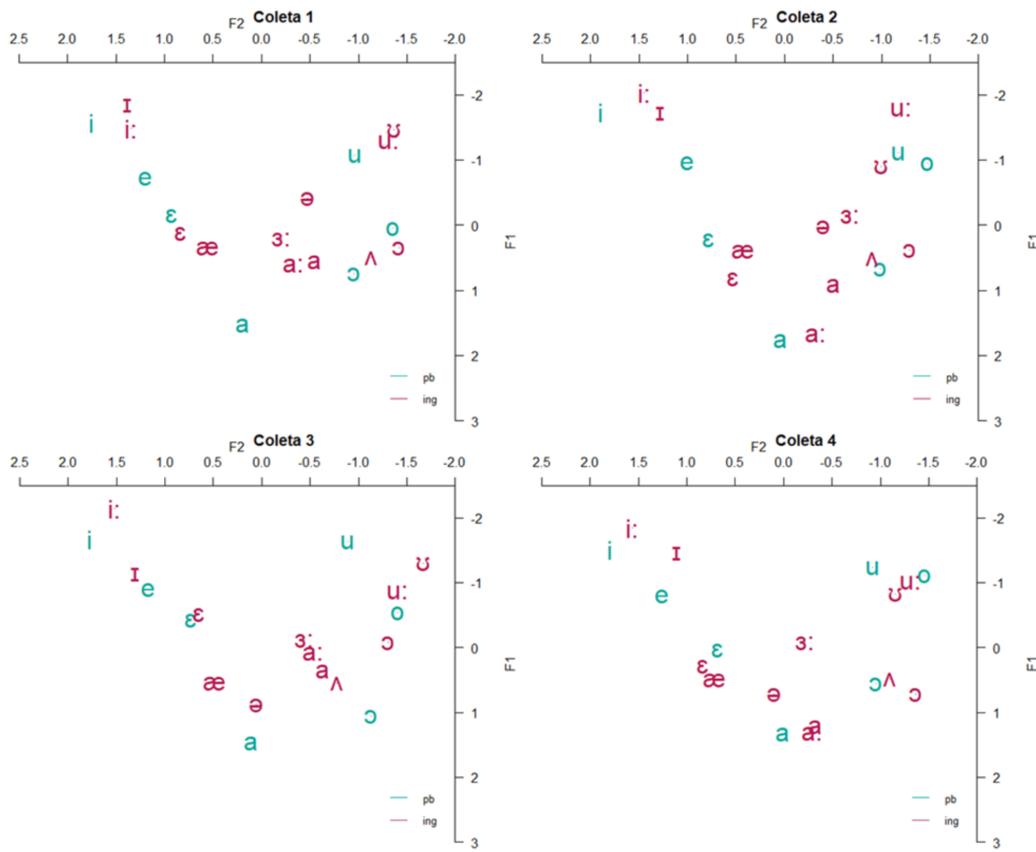


Fonte: Elaborado pelos autores.

Figura 13 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 5.

Notam-se alguns efeitos do retorno à L1 no desenvolvimento desse participante. Houve um abaixamento na vogal /i/ do PB durante o período de imersão (coletas 2 e 3), mas que retornou ao estado inicial após o retorno à L1 (coleta 4). Também se percebe um distanciamento maior das vogais /e/ e /ε/ na última coleta. A imersão na L2 ocasionou um alçamento da vogal /u/, que se manteve até o retorno ao Brasil, causando um distanciamento maior das vogais /u/ e /ɔ/ em relação ao estado inicial (coleta 1). Em relação à L2, o primeiro ponto de imersão (coleta 2) proporcionou uma perda na distinção entre pares de vogais longas e curtas na parte central e baixa do espaço acústico, mas a distinção dessas vogais pela duração foi reestabelecida nas próximas coletas. Mesmo após a imersão, o participante teve dificuldade em distinguir vogais centrais e o par /u/ e /ʊ/.

O Participante 6 é natural do Rio Grande do Sul, tinha 31 anos no início da pesquisa e residiu no estado de Rhode Island. Seu índice de imersão aumentou consideravelmente ao longo das coletas. Esse participante reportou atividades quase diárias no laboratório da universidade, usando o inglês, viagens e participação em eventos acadêmicos, que proporcionaram o uso do inglês, além de pouca comunicação com falantes de português.

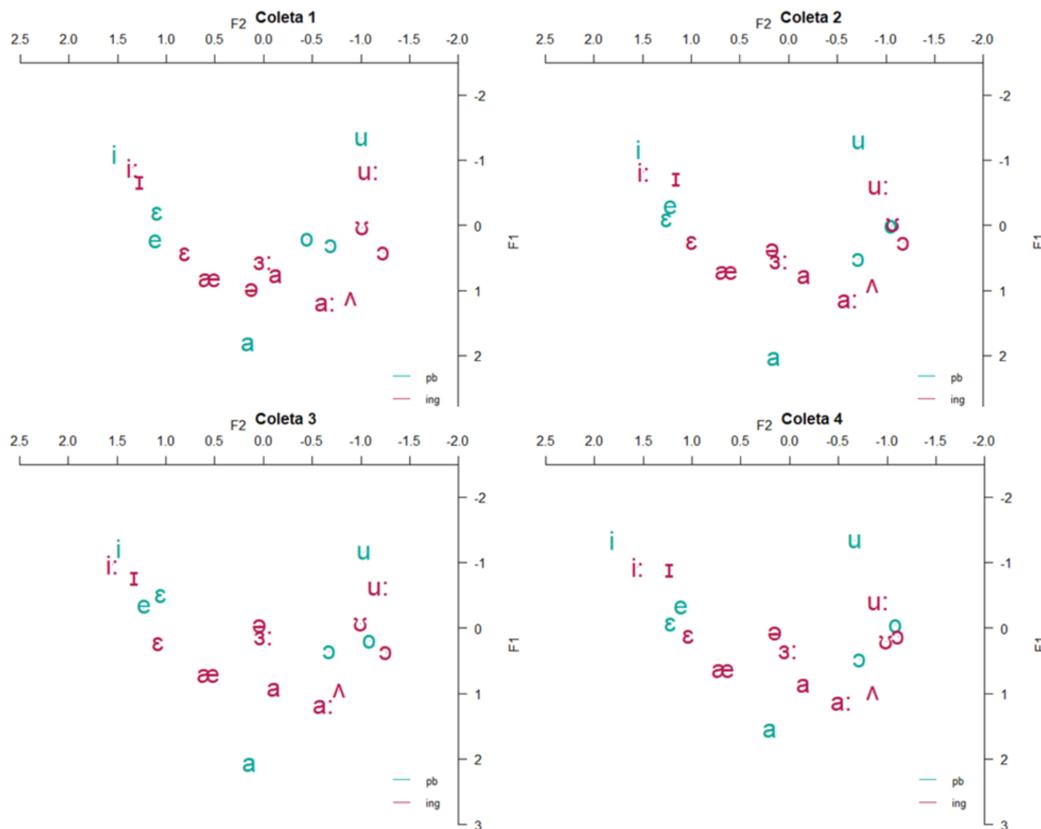


Fonte: Elaborado pelos autores.

Figura 14 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 6.

A imersão na L2 durante a coleta 2 ocasionou um distanciamento entre as vogais /e/ e /ɛ/ do PB, mas isso foi revertido e o distanciamento entre essas vogais voltou como era no estado inicial nas coletas seguintes. A imersão no período da coleta 2 também proporcionou um abaixamento da vogal /a/ e um alçamento da vogal /o/ em direção à vogal /u/ do PB e à vogal /ʊ/ do inglês. No IL2, percebe-se uma variação na produção da vogal /a:/, a qual iniciou em posição central, abaixou no primeiro mês de imersão, retornou à posição central no nono mês e, no retorno ao Brasil, finalizou, em altura, próxima da vogal /a/ do PB, causando uma reorganização no sistema. Esse participante também foi capaz de distinguir os pares /i:/ - /ɪ/, /ɛ/ - /æ/ e /u:/ - /ʊ/ até a coleta 3, mas não manteve a distinção do último par no retorno ao Brasil. Também se nota uma alta variação na produção das vogais na região central dos gráficos, demonstrando que o participante ainda tem dificuldade em estabelecer essas categorias no espaço acústico.

O Participante 7 é natural de Minas Gerais, tinha 28 anos no início da pesquisa e residiu na Flórida. Seu índice de imersão foi o único que reduziu entre a coleta 1 e a 2. O participante reportou que foi recebido por brasileiros logo na sua chegada, pois morou em uma comunidade de brasileiros com seu companheiro. Apesar de um contexto social que favoreça uma menor imersão na L2, o participante frequentou a universidade diariamente e interagiu com falantes de inglês L1 constantemente.



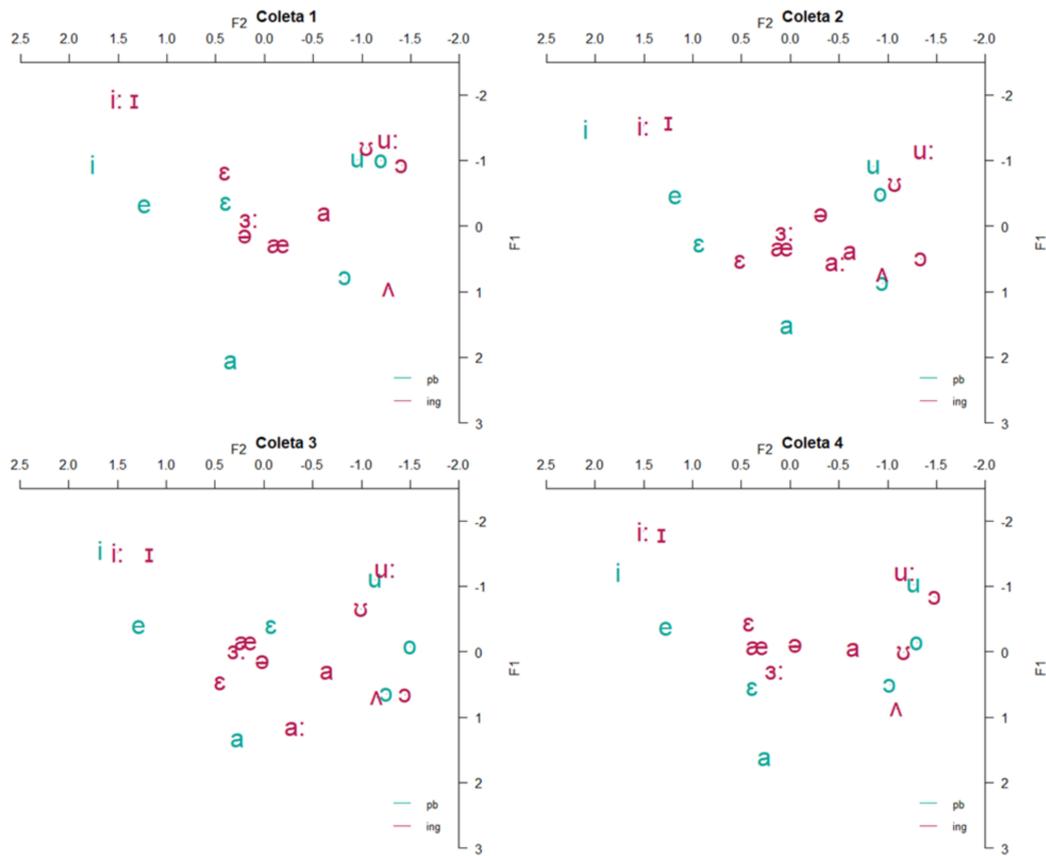
Fonte: Elaborado pelos autores.

Figura 15 - Distribuição das vogais do PB e do IL2, por coleta, do participante 7.

Em relação ao sistema da L1, percebe-se que há uma variação na posição das vogais /ɛ/ e /e/ na primeira (antes da imersão) e terceira (nono mês de imersão) coletas, além de uma movimentação de /o/ em todos os pontos. Acreditamos que houve desvios de produção de /ɛ/ e /o/ durante os experimentos e que não seja reflexo da L2. Ainda no PB, houve um alçamento da vogal /a/ e uma anteriorização da vogal /i/ após a imersão.

A imersão no IL2 favoreceu uma grande movimentação das vogais. O participante foi capaz de aumentar a distinção dos pares /i/ - /ɪ/ e /ɛ/ - /æ/ a partir da segunda coleta, mantendo a distinção até o retorno ao país de sua L1. Como estratégia para diferenciar /u:/ de /ʊ/, aproximou a segunda da categoria da vogal /o/ na L1, enquanto a primeira movimentou-se bastante no espaço acústico, mas nunca fazendo sobreposição. Nas duas primeiras coletas, nota-se uma semelhança entre as vogais centrais, as quais distanciaram-se no retorno ao país de sua L1, com um abaixamento de /a/.

O Participante 8 é natural de São Paulo, residiu na Califórnia e tinha 31 anos durante as coletas. Esse participante reportou problemas de moradia nos primeiros meses do intercâmbio, residiu com um brasileiro e realizou suas atividades remotamente, o que não contribuiu para um aumento no índice de imersão na segunda coleta. Porém, houve um aumento significativo na coleta 3. Ele mencionou que, no período dessa coleta, participava das atividades na universidade quase diariamente, fazendo uso do inglês, e que não dividia mais moradia com brasileiro.

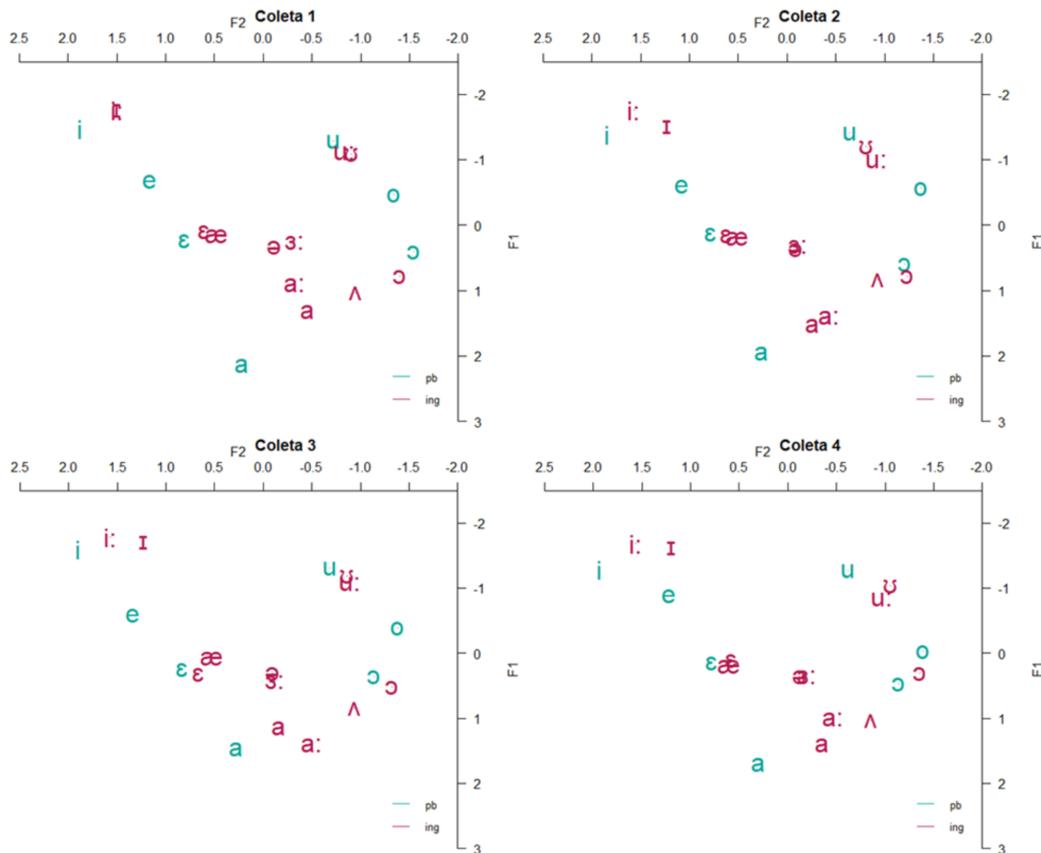


Fonte: Elaborado pelos autores.

Figura 16 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 8.

A produção das vogais no PB variou ao longo das quatro coletas, ocasionando médias em espaço acústico não esperados. A vogal /i/ foi produzida com maior anterioridade na segunda coleta, enquanto a vogal /ε/ foi centralizada nas coletas 1 e 3. A vogal /a/ foi alçada desde a segunda coleta, permanecendo nesta posição até o retorno ao Brasil, assim como a vogal [o] foi produzida mais baixa. No IL2, percebe-se que o participante teve uma tendência a centralizar as vogais médias e baixas, produzindo-as com categorias semelhantes, mesmo durante imersão na L1.

O Participante 9 é natural de Alagoas, tinha 29 anos e residiu em Nebraska com companheiro brasileiro. O índice de imersão na primeira coleta foi baixo em comparação aos outros participantes, mas teve um aumento significativo na coleta 2, quando o participante iniciou as atividades na universidade. Porém, o índice caiu na terceira imersão, quando reportou não estar mais indo à universidade por ter finalizado sua pesquisa. Também reportou contato frequente com brasileiros na coleta 3.



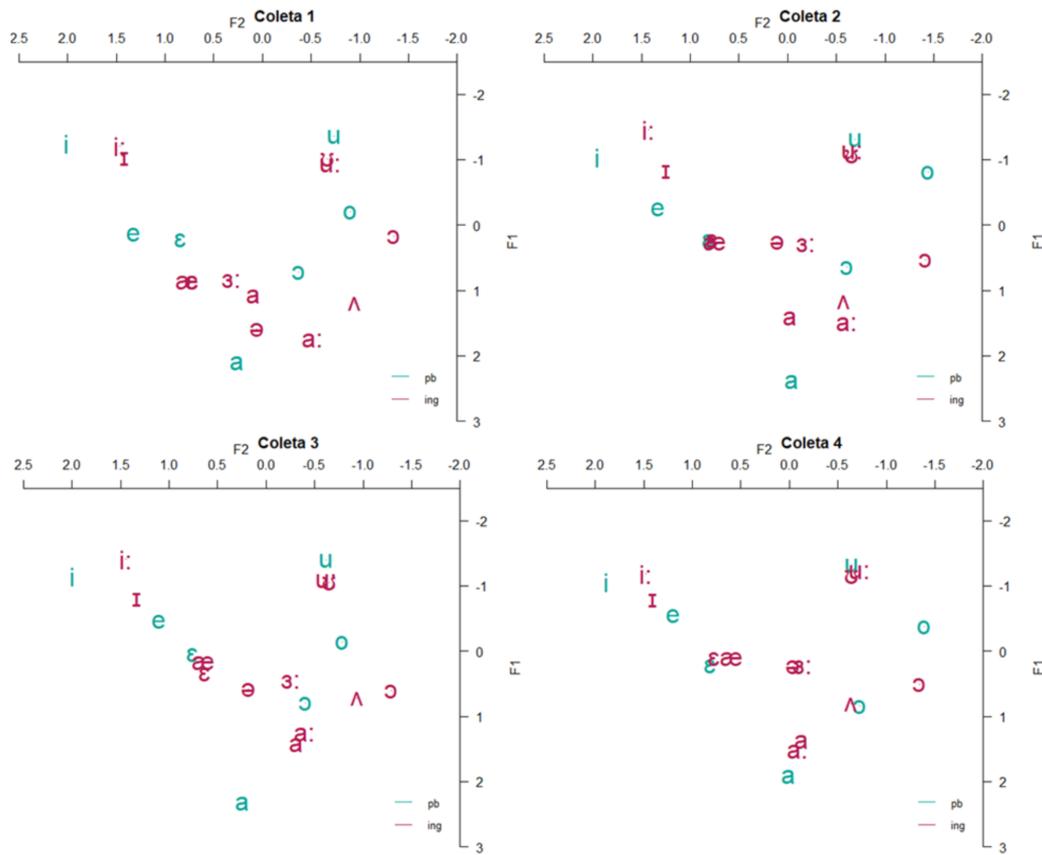
Fonte: Elaborado pelos autores.

Figura 17 - Distribuição das vogais do PB e do IL2, por coleta, do Participante 9.

No PB, a vogal /ɔ/ foi produzida com maior anterioridade desde a segunda coleta, permanecendo até o retorno ao Brasil. Na terceira coleta, a vogal /i/ aproximou-se da equivalente no inglês, mas retornou ao estado inicial na coleta 4. A partir da coleta 3, a vogal /a/ também foi produzida mais alta do que na coleta 1, atingindo maior proximidade com a equivalente na L2 no nono mês de imersão.

No IL2, o participante teve dificuldade em distinguir pares problemáticos na primeira coleta, produzindo as vogais, aparentemente, na mesma categoria. Após o retorno ao país da L1, foi capaz de distinguir satisfatoriamente /i:/ e /ɪ/, porém, não teve êxito nos pares /æ/-/ε/ e /u:/-/ʊ/ associados as vogais /ε/ e /u/ do PB, respectivamente. Também teve dificuldade em distinguir as vogais /ə/ e /ɜ:/ a partir da segunda coleta, embora já fizesse essa distinção inicialmente.

O Participante 10 é natural de Mato Grosso, tinha 37 anos durante a pesquisa e residiu no Texas. No momento da coleta 2, o participante tinha iniciado suas atividades na universidade, estava conhecendo a cidade em que residia e dividia moradia com outra falante de IL2, contribuindo para um aumento considerável no índice de imersão. Porém, na coleta 3, o participante estava finalizando sua pesquisa de maneira remota há alguns meses, não estava indo à universidade e estava morando com seu companheiro brasileiro, o que resultou em uma queda no seu nível de imersão.



Fonte: Elaborado pelos autores.

Figura 18 - Distribuição das vogais do PB e do IL2, por coleta, do participante 10.

No PB, o único movimento considerável no espaço acústico é o da vogal /o/, variando sua produção no eixo horizontal. A vogal /a/ aproximou-se das equivalentes do IL2 após o retorno ao PB. No IL2, o participante dominou a distinção no par /i:/ e /ɪ/ desde a segunda coleta. Porém, não conseguiu criar novas categorias no IL2 para as vogais próximas de /ɛ/ e /u/ do PB, mesmo no ápice da imersão. As vogais centrais do IL2 também apresentaram sobreposição, não sendo distintas satisfatoriamente as vogais longas das curtas no nível formântico.

#### 4 DISCUSSÃO

De maneira geral, os efeitos da imersão no IL2 no sistema vocálico do PB de brasileiros foram tímidos. Embora tenham sido observados movimentos nas categorias vocálicas do PB no nível intraindividual, os movimentos mais evidentes e significativos foram em relação ao desenvolvimento de categorias vocálicas do IL2 (a direção comumente estudada). Porém, algumas constatações de influências mútuas entre as duas línguas podem ser discutidas no desenvolvimento de cada participante.

Por exemplo, constataram-se movimentos de vogais do PB em direção à formação de uma categoria única nas proximidades das vogais equivalentes do IL2 (como /i/ do PB em direção à /i:/ e /ɪ/ do IL2, Participante 8). Essa é uma movimentação inversa, porém possível (Schmid; Kopke, 2019), da comumente esperada, na qual a vogal da L1 atua como atratora para as vogais da L2 (Beckner et

al., 2009). Também se percebeu que o ápice da imersão na L2 com pouco uso da L1 pode influenciar na distribuição vocálica no espaço acústico (Linck; Kroll, 2019), como no caso do abaixamento de /a/ do Participante 8. Embora falantes altamente imersos na L2 sejam os candidatos ideais para a emergência de atrito linguístico (Schmid, 2011, 2019), o Participante 3, com os maiores índices de imersão, não apresentou tanta variação nas vogais de sua L1 quanto outros participantes com índices menores. Entretanto, constataram-se mudanças na distribuição vocálica do PB, mesmo após o retorno ao Brasil, mas não é possível saber se esses efeitos foram duradouros ou apenas temporários (Chang, 2019).

O desenvolvimento do IL2 foi mais perceptível, mostrando que a imersão favoreceu uma reorganização das vogais do inglês. De maneira geral, todos os participantes beneficiaram-se do input recebido durante o intercâmbio. Alguns participantes foram capazes de criar novas categorias para pares de vogais problemáticos e superar a influência de vogais atratoras do PB (como os participantes 9 e 10), mesmo com índices de imersão relativamente baixos. Isso indica que a imersão na L2 é um fator que favorece a aprendizagem de categorias vocálicas do inglês, as quais aparentam ser mais difíceis de dominar quando em imersão na L1 ou sem instrução prévia.

Também observamos períodos de alta variabilidade e reorganização do sistema vocálico, em que falantes foram capazes de distinguir vogais semelhantes, porém, perdendo a distinção entre outros pares previamente dominados. Para alguns falantes, esse período de perda de características já desenvolvidas foi temporário, retornando seu sistema ao estado inicial após a reimersão na L1. Para outros, a distinção entre as vogais continuou presente após o retorno ao Brasil. Ambos os casos são exemplos da dinamicidade e não linearidade inerentes ao sistema linguístico (Beckner et al., 2009).

## 5 CONSIDERAÇÕES FINAIS

Damos como confirmada a hipótese levantada para a pergunta-problema proposta, pois os maiores efeitos da imersão na L2 foram constatados no desenvolvimento da própria L2 e apenas vestígios de mudanças na distribuição vocálica da L1 foram observados. O sistema vocálico da língua inglesa foi beneficiado pela imersão em um contexto na qual ela é a predominante, resultando em uma reorganização positiva de categorias vocálicas maior do que antes do período de imersão. Por outro lado, pouca reorganização foi percebida nas categorias vocálicas da L1, assemelhando-se ao estado inicial, antes da imersão. Esses resultados reforçam que influências mútuas entre as duas línguas são possíveis, porém, são menos frequentes no sentido L2 – L1.

O objetivo geral delimitado foi alcançado. A imersão na L2 mostrou-se um fator importante para o desenvolvimento do sistema da L2, como discutido anteriormente. Porém, é importante reforçar a existência de percursos distintos no nível intraindividual, sendo possível perceber diferentes tipos de mudança: um novo estado do sistema após à reimersão na L1, bem como desvios temporários. Essa não linearidade também foi observada na L1, a qual proporcionou alguns casos de reorganização no PB, mas, no geral, não foi um fator causador de mudanças.

Como consideração final, acreditamos que uma análise com mais pontos de coletas durante a imersão seria fundamental para analisar o desenvolvimento do

sistema vocálico dos participantes, bem como uma coleta posterior ao retorno ao Brasil. Embora esta análise tenha revelado mudanças no sistema vocálico dos participantes na última coleta (no retorno ao Brasil), uma nova coleta, meses depois, seria importante para investigar a durabilidade desses efeitos. Além disso, mais pontos de coleta durante a imersão em si seriam valiosos para entender o percurso de desenvolvimento do sistema estudado. Infelizmente, coletas em outros pontos planejados não foram possíveis devido a dificuldades pessoais durante o período desta pesquisa.

Também esperamos analisar a duração das vogais no nível intraindividual, pois pode ser uma estratégia de distinção vocálica adotada pelos participantes. É possível que alguns participantes dominem a distinção de certas vogais do inglês no nível formântico, mas não as diferencie satisfatoriamente através da duração. Também é possível que a duração seja a estratégia de distinção adotada por algum participante, mesmo não dominando essa distinção no nível formântico. Essa análise não foi realizada em conjunto com o nível formântico devido à limitação na extensão da pesquisa atual para publicação.

Por fim, o ajuste de modelos estatísticos para analisar a distância de pares vocálicos do IL2 se faz necessário para a continuidade deste estudo. Assim, será possível mensurar o efeito da imersão no desenvolvimento do sistema vocálico dos participantes, bem como averiguar, apropriadamente, se a movimentação das categorias vocálicas analisadas ao longo das coletas foi significativa.

## REFERÊNCIAS

Anderson J, et al. The language and social background questionnaire: assessing degree of bilingualism in a diverse population. *Behavior Research*. 2018;50(1):250-263.

Beckner C, et al. Language is a complex adaptive system: position paper. *Language Learning*. 2009;59(1):1-26.

Boersma P, Weenink D. Praat: doing phonetics by computer [programa de computador]. Version 6.1.03. Amsterdam: Universiteit van Amsterdam; 2021. Disponível em: <http://www.fon.hum.uva.nl/praat>.

Boersma P, Weenink D. Praat: doing phonetics by computer [programa de computador]. Version 6.3.10. Amsterdam: Universiteit van Amsterdam; 2023. Disponível em: <http://www.fon.hum.uva.nl/praat>.

Chang C. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of Phonetics*. 2012;40(2):249-268.

Chang C. A novelty effect in phonetic drift of the native language. *Journal of Phonetics*. 2013;4(6):520-533.

Chang C. Phonetic drift. In: Schmid M, Kopke B, editores. *The Oxford handbook of language attrition*. Oxford: Oxford University Press; 2019. p. 191-203.

De Bot K. Dynamic systems theory, life span development and language attrition. In: Kopke B, et al., editores. *Language attrition: theoretical perspectives*. Amsterdam: John Benjamins Publishing Company; 2007. p. 53-68.

Kartushina N, Martin C. Third-language learning affects bilinguals' production in both their native languages: a longitudinal study of dynamic changes in L1, L2 and L3 vowel production. *Journal of Phonetics*. 2019;77:100920.

Lang B, Davidson L. Effects of exposure and vowel space distribution on phonetic drift: evidence from American English learners of French. *Language and Speech*. 2019;62(1):30-60.

- Larsen-Freeman D. Chaos/complexity science and second language acquisition. *Applied Linguistics*. 1997;18(2):141-165.
- Leeuw E, Opitz C, Lubinska D. Dynamics of first language attrition across the lifespan. *International Journal of Bilingualism*. 2013;17(6):667-674.
- Lima Jr. R. Análise longitudinal de vogais do inglês-L2 de brasileiros. *Gradus: Revista Brasileira de Fonologia de Laboratório*. 2016;1(1):145-176.
- Linck J, Kroll J. Memory retrieval and language attrition: language loss or manifestations of a dynamic system? In: Schmid M, Kopke B, editores. *The Oxford handbook of language attrition*. Oxford: Oxford University Press; 2019. p. 88-97.
- Mccloy D. phonR: tools for phoneticians and phonologists. R package version 1. 0-7 [programa de computador]. 2022. Disponível em: <https://CRAN.R-project.org/package=phonR>.
- Opitz C. A complex dynamic systems perspective on personal background variables in L attrition. In: Schmid M, Kopke B, editores. *The Oxford handbook of language attrition*. Oxford: Oxford University Press; 2019. p. 49-62.
- R Core Team. R: a language and environment for Statistical Computing [programa de computador]. Version 4.3.2. Viena: R Foundation for Statistical Computing; 2023. Disponível em <http://www.R-project.org/>.
- Schereschewesky L, Alves U, Kupske F. First language attrition: the effects of English (L2) on Brazilian Portuguese VoT patterns in an L1-dominant environment. *Letrônica*. 2011;10(2):700-716.
- Schmid M. *Language attrition*. Cambridge: Cambridge University Press; 2011.
- Schmid M, Kopke B. *The Oxford handbook of language attrition*. Oxford: Oxford University Press; 2019.

# Diferenças na produção de foco prosódico contrastivo na fala de adultos e de crianças com aquisição fonológica típica e atípica do Português Brasileiro

## *Differences in contrastive focus marking in speech of adults and children under typical and atypical phonological acquisition of Brazilian Portuguese*

Geovana Soncin\*

*Universidade Estadual Paulista (UNESP), Marília, SP, Brasil*

Cecília Lorena Silva Guida\*\*

*Universidade Estadual Paulista (UNESP), Marília, SP, Brasil*

Fernanda Leitão de Castro Nunes de Lima\*\*\*

*Universidade Estadual Paulista (UNESP), Marília, SP, Brasil*

Larissa Cristina Berti\*\*\*\*

*Universidade Estadual Paulista (UNESP), Marília, SP, Brasil*

**Resumo:** O presente artigo tem como objetivo geral caracterizar acusticamente a marcação de foco prosódico na fala de adultos, de crianças com desenvolvimento típico de linguagem e de crianças com transtorno fonológico a fim de verificar como esses grupos de sujeitos marcariam o foco. Amostras de fala foram obtidas a partir de experimento elaborado para elicitare sentenças com foco prosódico contrastivos e sentenças neutras. Análise acústica de frequência fundamental, duração e intensidade foi realizada nas sentenças que compuseram as amostras de fala dos três grupos. O estudo mostra performance distinta dos grupos na marcação de foco contrastivo na análise dos três parâmetros fonéticos e sugere que crianças com transtorno fonológico podem apresentar instabilidades no plano prosódico.

**Palavras-chave:** Aquisição fonológica. Produção de fala. Prosódia. Foco. Transtorno fonológico.

**Abstract:** The goal of this article is to characterize acoustically and compare prosodic focus marking in speech of adults, children under typical language development and children diagnosed with phonological disorders. Speech samples were obtained from an experiment designed to elicit sentences with contrastive focus and neutral sentences. Acoustical analysis of fundamental frequency, duration and intensity was performed on the sentences produced by each participant's groups. The results show that groups performed differently in marking contrastive focus by using the three phonetic parameters and suggest that children with phonological disorders may present instabilities in the prosodic plan.

---

\* Professora do Programa de Pós-Graduação em Fonoaudiologia e bolsista FAPESP Jovem Pesquisador (Proc. 2020/10144-3), Universidade Estadual Paulista, Marília, SP, Brasil; [geovana.soncin@unesp.br](mailto:geovana.soncin@unesp.br)

\*\* Mestranda do Programa de Pós-Graduação em Fonoaudiologia, Universidade Estadual Paulista, Marília, SP, Brasil; [cecilia.lorena@unesp.br](mailto:cecilia.lorena@unesp.br)

\*\*\* Doutoranda do Programa de Pós-Graduação em Fonoaudiologia, Universidade Estadual Paulista, Marília, SP, Brasil; [fernanda.leitao@unesp.br](mailto:fernanda.leitao@unesp.br)

\*\*\*\* Professora do Departamento de Fonoaudiologia, Universidade Estadual Paulista, Marília, SP, Brasil; [larissa.berti@unesp.br](mailto:larissa.berti@unesp.br)

**Keywords:** Phonological acquisition. Speech production. Prosody. Focus. Phonological disorder.

## 1 INTRODUÇÃO

Este trabalho caracteriza-se como um estudo descritivo sob os pressupostos da análise linguística de base fonético-fonológica cujos resultados podem reverberar na área dos estudos clínicos fonoaudiológicos com ênfase em aquisição de linguagem. Nessa área, a consideração de aspectos prosódicos com base em critérios linguísticos para a elucidação de diagnóstico e casos clínicos é pouco frequente (Diehl; Paul, 2009). A esse respeito, vale mencionar publicação recente (Hawthorne; Fischer, 2020) que, como resultado de entrevistas realizadas com 245 fonoaudiólogos, mostrou que, embora esses profissionais reconheçam que a prosódia esteja no âmbito de sua atuação, raramente ela é avaliada ou tratada em suas práticas privilegiadamente por dois motivos: (i) falta de conhecimento sobre a natureza da prosódia; e (ii) desconhecimento de métodos eficazes para avaliação e tratamento da prosódia.

Particularmente, no campo da aquisição de linguagem, a compreensão sobre como se daria o desenvolvimento prosódico é lacunar e pode ser considerada um desafio para a fonoaudiologia (Peppé, 2009). Nesse cenário, estudos que caracterizem prosodicamente a produção de fala de sujeitos típicos, adultos e crianças, fazem-se necessários para a obtenção de dados de referência a partir dos quais possam ser comparados dados de fala de populações que potencialmente seriam afetadas por processos de aquisição prosódica desviante.

Inserido nesse contexto de demandas sobre melhor entendimento sobre características prosódicas manifestadas na fala de crianças, este trabalho analisa acusticamente a produção de foco prosódico contrastivo na fala de adultos, de crianças em desenvolvimento típico de linguagem (DTL) e de crianças diagnosticadas com transtorno fonológico (TF). Partindo da hipótese de que a fala de adultos e crianças apresentam diferenças que correspondem à condição de desenvolvimento em que se encontram – embora possam ser equiparavelmente competentes comunicativamente consideradas suas condições de desenvolvimento –, tem-se como objetivo descrever, em primeira instância, quais parâmetros acústicos marcam a focalização prosódica na fala desses grupos e, em segunda instância, identificar possíveis diferenças e semelhanças entre eles.

A proposta de investigação do foco prosódico em amostras de fala de crianças com TF em comparação a crianças em DTL e adultos justifica-se por duas razões. Por um lado, a literatura fonoaudiológica com ênfase em aquisição de fala reporta erros prosódicos relativos à produção de acento lexical, acento frasal e foco como características discriminativas dos transtornos motores de fala, especialmente a *apraxia de fala na infância* (Shriberg; Aram; Kwaitkowski, 1997; Shriberg et al., 2019). Destacamos, porém, que a caracterização desses erros com base em critérios de natureza linguística não é apresentada por esses trabalhos. Por outro lado, em estudo anterior de nossa autoria (Berti et al., in preparation) observamos que crianças com TF que apresentam o subtipo *atraso no desenvolvimento fonológico* foram diferenciadas em dois grupos devido ao desempenho prosódico que apresentaram em tarefas envolvendo aspectos prosódicos da produção da fala, quais sejam: acento lexical e foco.

Esse último resultado particularmente nos motivou a aprofundar a investigação sobre o desempenho prosódico em crianças com TF uma vez que, em termos de impacto, pode sugerir que alterações no âmbito prosódico possam afetar não apenas pacientes com quadro clínico de natureza motora, como a apraxia de fala, mas também pacientes com transtornos de fala que se classificam no plano fonológico. Assim, a fim de melhor compreender como se caracterizariam contrastes prosódicos na fala de crianças com TF, elegeram-se, neste trabalho, o *foco* como categoria de análise nas amostras de fala dessas crianças pelo fato de, nos estudos anteriormente reportados, o *foco* ter sido apontado como categoria prosódica em que se podem identificar performances distintas das populações com transtorno motor ou fonológico. Para tanto, porém, torna-se necessário tomar sujeitos típicos, sejam eles adultos e/ou crianças, como pontos de referência de modo a identificar pontos de congruência e/ou divergência pela comparação entre suas amostras.

Conforme nomenclatura da *American Speech-Language-Hearing Association* (ASHA, 2023), o TF enquadra-se nos chamados *transtornos dos sons da fala*, os quais podem estar relacionados a dificuldades de percepção e produção de fala, representação fonológica e/ou planejamento motor, ocasionando alterações de produção de fala que afetam a inteligibilidade da produção linguística e, conseqüentemente, a qualidade da comunicação verbal. Especificamente, de acordo com o *Manual Diagnóstico e Estatístico de Transtornos Mentais – 5ª. Edição* (DSM-5; *American Psychiatric Association*, 2014), o TF – transtorno com o qual as crianças de um dos grupos de sujeitos contemplados neste trabalho foram diagnosticadas – se caracteriza por uma dificuldade persistente na produção da fala, que se manifesta na forma de erros definidos por inconsistências na representação de um sistema fonológico em aquisição. O diagnóstico do TF ocorre quando a fala apresentada pela criança está aquém do esperado para sua idade e estágio de desenvolvimento.

Dodd (2005) apresenta uma classificação linguística que identifica subtipos de TF com base em diferentes características linguísticas observadas nos erros apresentados pelas crianças na produção de fala. Dentre eles estão: *atraso no desenvolvimento fonológico*, *transtorno fonológico atípico consistente* e *transtorno fonológico inconsistente*. Apesar das suas diferenças, que podem ser relevantes para o diagnóstico diferencial e para a tomada de decisão referente ao planejamento terapêutico de cada paciente, em todos os casos, o sistema fonológico em aquisição pela criança encontra-se com problemas, na medida em que neutraliza contrastes fonêmicos e/ou estruturas silábicas que compõem a representação linguística da língua em aquisição.

Em termos de produção de fala, a caracterização segmental da produção linguística das crianças com TF é largamente descrita na literatura (Wertzner, 2003; Patah; Takiuchi, 2008; Hearnshaw; Baker; Munro, 2018; Berti et al., 2020). No que diz respeito à caracterização prosódica, porém, o mesmo não ocorre. No entanto, tem sido levantadas problematizações na literatura acerca da possibilidade de a organização prosódica da fala das crianças com TF apresentar inconsistências e/ou diferenças em relação às crianças em desenvolvimento típico de mesma faixa etária.

Polli (2022) comparou a performance de crianças em DTL e de crianças com TF em três tarefas que envolviam contrastes prosódicos, tais como acento de palavra, foco prosódico e marcação de fronteira final e não final, e mostrou, como resultados, que o grupo de crianças com TF apresentou performance inferior na marcação desses

contrastes tanto por meio de pistas verbais quanto por meio de pistas visuais. Por sua vez, o estudo de Soncin, Polli e Berti (2022) mostrou que, do ponto de vista acústico, a marcação de foco prosódico contrastivo na fala de crianças com TF, embora possa apresentar aumento de duração e intensidade como esperado para o português brasileiro, não é marcada pela configuração tonal característica da marcação do foco contrastivo nessa língua.

Desse modo, considerando os indícios que esses estudos têm apresentado, bem como tomando como pressuposto que a prosódia é parte do sistema fonológico de uma língua (Nespor; Vogel, 1986, 2007), o presente trabalho analisa amostras de fala de crianças com TF para caracterizar aspectos prosódicos – notadamente por meio da diferenciação entre sentenças neutras e focalizadas – a fim de identificar se haveria diferenças em relação a crianças em DTL. Não obstante, supondo que, do ponto de vista aquisicional, crianças podem apresentar diferenças em relação a adultos, o presente trabalho contempla os três grupos de sujeitos. A contextualização teórica e descritiva acerca da focalização prosódica, bem como o estado da arte dos estudos sobre aquisição da prosódia que contemplaram o foco prosódico, são apresentados na próxima sessão.

## 2 PRESSUPOSTOS TEÓRICOS

Ao se pensar na aquisição de linguagem de uma criança, consideram-se diversos aspectos que coocorrem ao longo do processo, tais como: aspectos sintáticos, lexicais, morfológicos, fonológicos. Para a caracterização de como esses aspectos são adquiridos, é preciso organizá-los sistematicamente de acordo com o funcionamento que assumem na língua materna alvo e, como tal, qual o seu efeito para o objetivo de uma comunicação eficiente (De Ruiter, 2012).

O aspecto fonológico se destaca por ser objeto de inúmeros estudos, principalmente com intuito de descrever sua aquisição típica e, por conseguinte, atípica. A literatura descreve quais são os processos fonológicos típicos para cada faixa etária, considerando as classes fonêmicas e a complexidade silábica que compõem a fonotática do português brasileiro (PB), a partir de resultados obtidos pela aplicação de instrumentos de avaliação fonológica (Yavas, 1988; Lamprecht, 2004; Berti; Pagliuso; Lacava, 2009; Ceron et al., 2020). Assim, a partir da descrição da aquisição fonológica de crianças em DTL, é possível definir padrões atípicos de aquisição e ainda criar métodos de avaliação e terapia para esses (Gonçalves; Keske-Soares; Checalin, 2010).

No entanto, conforme é bem estabelecido nos estudos linguísticos, no plano fonológico, incluem-se também os aspectos prosódicos que caracterizam a língua materna de uma criança. Portanto, à fonologia dizem respeito não apenas os aspectos segmentais, mas também a organização prosódica. Essa visão, adotada no presente trabalho, está ancorada no modelo de Fonologia Prosódica defendido por Nespor e Vogel (1986, 2007), segundo o qual a prosódia atua no sistema linguístico fonológico de um falante como estrutura organizadora da fala. Considerada desse modo, assume-se a existência uma hierarquia prosódica subjacente à língua composta por constituintes prosódicos, os quais são delimitados por fronteiras prosódicas e se caracterizam por apresentar proeminências nos vários níveis (Nespor; Vogel, 1986, 2007).

A adoção de um modelo fonológico para a descrição prosódica conforme adotado neste trabalho justifica-se, entre outras razões, pois através dele é possível prever fronteiras prosódicas e pontos proeminentes na fala de modo tal que, numa análise acústica, a variação dos parâmetros fonético-acústicos seriam efeitos da estrutura prosódica e, como tal, se manifestariam de maneira relativamente sistemática numa dada língua. Assim, variações de frequência fundamental, duração e intensidade são analisadas sob a ótica que considera que eles ocorrem em pontos previstos pela estrutura prosódica.

Para além da visão estrutural formal da prosódia nos termos propostos por Nespor e Vogel (1986, 2007), sabe-se que, no plano comunicativo, a prosódia desempenha diferentes funções (Ladd; Cutler, 1983; Geluykens; Swerts, 1994). No interior do amplo espectro de funções prosódicas desempenhadas na comunicação, há unanimidade, entre diversos estudiosos, quanto à marcação de proeminências ser uma das funções linguísticas fundamentais desempenhadas pela prosódia (Terken; Hermes, 2000; D’Imperio et al., 2005; Gussenhoven, 2006). Em termos gerais, a marcação de proeminência caracteriza-se como um meio de distinguir informações importantes daquelas que são menos importantes para os fins comunicativos de uma dada interação verbal. Tal fenômeno ocorre pois os falantes de uma língua variam as características prosódicas de seus enunciados da fala em resposta a mudanças da estrutura informacional a eles subjacentes.

Na definição de Chafe (1976), estrutura informacional se refere à divisão de um enunciado em blocos de informação, com o objetivo de permitir uma atualização contínua sobre os pontos comuns de informação compartilhados pelos participantes da interação. De acordo com Krifka (2008), as categorias de *foco*, *tópico* e *informação dada* são essenciais para a descrição da estrutura informacional dos enunciados linguísticos com vistas ao seu caráter comunicativo. Dentre elas, a categoria de foco, abstrata cognitivamente, pode ser definida como uma indicação da “presença de alternativas de informação que são relevantes para a interpretação das expressões linguísticas” num dado contexto comunicativo (Krifka, 2008, p. 247, tradução nossa). Em termos de manifestação na materialidade linguística, o foco pode ser prosodicamente expresso com especificidades que variam entre as línguas do mundo.

Para ilustrar a definição que apresenta para a categoria de foco, Krifka (2008) utiliza o exemplo de um enunciado resposta à pergunta *Quem pegou o cookie?*, sendo ele: *PEDRO pegou o cookie*<sup>1</sup>. Nesse exemplo, a pergunta com o pronome interrogativo *quem* solicita a atualização do contexto compartilhado entre os interlocutores para a identificação do referente desse pronome. No enunciado resposta, *Pedro* é o referente selecionado como a alternativa mais relevante dentre aquelas disponíveis no contexto, sendo, assim, o elemento focalizado do enunciado, ou seja, o elemento identificado como foco. Em decorrência da focalização, em termos prosódicos, *Pedro* recebe um acento tonal típico de foco em línguas como o inglês, língua do exemplo reportado, assim como pode ocorrer no português brasileiro conforme detalhamento apresentado adiante.

Há, porém, outra possibilidade de uso da categoria foco nas situações comunicativas quando se observa que, além da indicação da alternativa relevante para

---

<sup>1</sup> O exemplo apresentado no texto original é oriundo da língua inglesa e os enunciados do exemplo foram aqui por nós adaptados para o português do Brasil. As responsabilidades por essa adaptação são todas nossas.

interpretação das expressões linguísticas num dado contexto, o foco pode corrigir ou confirmar a informação que um interlocutor lança na interação verbal. Antes da explanação desse último tipo, vale nomeá-los: enquanto o primeiro tipo, cujo exemplo foi apresentado no parágrafo anterior, é chamado, por Krifka (2008), de foco estreito, o segundo é chamado de foco contrastivo, do qual tratamos a seguir.

Continuando com o exemplo da questão *Quem pegou o cookie?*, Krifka (2008) apresenta que um interlocutor B pode responder *A Maria pegou o cookie*, enquanto, imediatamente após essa afirmação, um interlocutor C retifica essa informação produzindo o enunciado *(Não), o PEDRO pegou o cookie*. Nesse enunciado produzido por C, o foco em PEDRO corrige a informação dada pelo interlocutor B, inserindo uma nova informação no contexto compartilhado entre os interlocutores, informação essa que contrasta – ou seja, que é distinta por oposição – com a informação anteriormente dada por B. No caso de um exemplo de confirmação da informação da dada pelo interlocutor B, Krifka (2008) explana que um interlocutor C' poderia proferir o enunciado *(Sim), a MARLA pegou o cookie*. Nos enunciados de C ou C', o elemento focalizado recebe acento tonal típico de focalização contrastiva, além de poder apresentar, a depender da língua, variações em outros parâmetros fonéticos (o detalhamento a esse respeito é apresentado a seguir).

Em termos de descrição prosódica das variedades do português, Frota et al. (2015) afirmam, baseando-se na tipologia apresentada por Krifka (2008), que as sentenças com foco estreito ou contrastivo se diferenciam das sentenças com foco largo (*broad focus*), uma vez que, enquanto as últimas são enunciadas sem a identificação de um único elemento como informação nova e, assim, são produzidas com entoação de sentenças declarativas neutras, as primeiras focalizam um elemento por meio de acentos tonais prototípicos, substituindo a interpretação de que toda a sentença traz informação nova pela 'novidade' interpretativa apresentada pelo elemento focalizado.

No português brasileiro, os dois tipos de foco foram identificados nas descrições apresentadas em Gonçalves (1997), Fernandes (2007), Moraes (2009), Carnaval, Moraes e Rilliard (2022); entre outros. Nesses trabalhos, porém, entende-se que ambos são do tipo estreito, sendo nomeados como *foco informacional* aquele que apresenta informação nova e *foco contrastivo* aquele responsável por corrigir o enunciado anterior e estabelecer contraste semântico com ele. Nesses trabalhos, ainda, o foco contrastivo se restringe à correção e não são descritos casos desse tipo de foco também com função confirmativa, como considerado por Krifka (2008).

Embora haja diferentes tipos de foco, o foco contrastivo é o tipo mais largamente estudado nos estudos de literatura prosódica em diferentes línguas, motivo que justifica sua seleção para o desenvolvimento do presente trabalho. Em línguas cujo sistema de marcação da estrutura informacional é baseado nas relações de proeminência (Kugler; Calhoun, 2020), como é o caso do português do Brasil, o padrão básico consiste no fato de a palavra focalizada em um enunciado tornar-se a palavra prosodicamente mais proeminente (Ladd, 2008). A proeminência é, então, marcada, por pistas fonéticas e fonológicas. Pistas fonéticas podem incluir valores mais altos de frequência fundamental (F0), aumento de duração e de intensidade e valores espectrais mais altos (Terken; Hermes, 2000; Ladd, 2008; Gussenhoven, 2006, 2011). Do ponto de vista fonológico, uma palavra é mais proeminente em um enunciado porque sua sílaba acentuada é elemento cabeça do constituinte prosódico mais alto do qual é parte (geralmente, a frase entoacional) e carrega, assim, o acento nuclear (Ladd, 2008). Nessa

perspectiva fonológica, conforme defende-se no Modelo de Fonologia Autossegmental e Métrica da Entoação (Pierrehumbert, 1980; Beckman; Pierrehumbert, 1980; Ladd, 2008; Arvaniti; Fletcher, 2020), o foco não é marcado de forma direta por pistas fonéticas, mas, ao contrário, de forma indireta: entende-se, pois, que as pistas marcam primariamente o acento nuclear, que, por sua vez, marca o foco (Ladd, 2008).

Neste trabalho, assumimos as descrições do português brasileiro realizadas por Fernandes (2007) e Frota et al. (2015) pelo fato de dialogarem com a visão de estrutura informacional proposta por Krifka (2008) e, especialmente, por serem descrições realizadas no quadro teórico da visão integrada entre Fonologia Prosódica (Nespor; Vogel, 1986, 2007) e Fonologia Autossegmental e Métrica da Entoação (Pierrehumbert, 1980; Beckman; Pierrehumbert, 1986; Ladd, 2008; Arvaniti; Fletcher, 2020).

Para a função de foco prosódico, tanto o acento tonal H\*+L quanto a combinação de acento tonal L\* + H seguido de tom de fronteira são reportados como configuração tonal atribuída à produção de foco nas variedades linguísticas do português do Brasil existentes na região Sudeste do país (Fernandes, 2007; Frota et al., 2015; Yano; Fernandes-Svartman, 2020). No plano da análise fonético-acústica, as pistas fonéticas que marcam o foco prosódico contrastivo no português do Brasil são: valores mais altos de frequência fundamental, aumento da intensidade e aumento da duração (Barbosa, 2012; Barbosa; Madureira, 2015). Em pesquisas com sujeitos adultos, para línguas como o inglês, descreve-se a frequência fundamental como a pista mais robusta para a marcação do foco, enquanto duração e intensidade seriam pistas secundárias (Terken; Hermes, 2000; Gussenhoven, 2006). Barbosa (2012) e Barbosa e Madureira (2015) afirmam que a mesma hierarquia de pistas é válida para os dados do português brasileiro.

Se, por um lado, a descrição da produção do foco na fala de adultos é amplamente descrita na literatura com base em critérios de natureza linguística, por outro, a descrição do foco na fala de crianças, se mantém lacunar, especialmente para o português brasileiro, ainda que estudos recentes tenham apresentado resultados para preencher essa lacuna, como veremos mais adiante.

No âmbito internacional do campo da pesquisa linguística com enfoque na aquisição de linguagem, estudos que propuseram avaliar o desenvolvimento prosódico em crianças em DTL tem procurado indicar, em que faixa etária média, determinadas habilidades prosódicas seriam adquiridas (Peppé; McCann, 2003; Wells; Peppé; Goulondris, 2004; Kalathottukaren; Purdy, 2017; Filipe et al., 2017) a partir da aplicação do instrumento de avaliação PEPS-C, sigla para *Profiling Elements of Prosodic Systems-Children* (Peppé; Mccann, 2003). A produção de foco prosódico contrastivo é uma das habilidades avaliadas por esses estudos, os quais indicam que a produção do foco contrastivo ocorreria após os nove anos de idade. No entanto, esses estudos, por estarem mais atentos ao desempenho das crianças do ponto de vista da performance comunicativa, não caracterizam como se manifestaria fonética e fonologicamente a marcação de foco prosódico na fala dessas crianças, nem mesmo se questionam se o modo de o marcar poderia ser distinto daquele realizado pelos adultos – cuja produção da fala é considerada o modelo linguístico a ser alcançado.

Por sua vez, o estudo de Wonnacott e Watson (2008), realizado no Reino Unido, que fez uso de análise acústica, relata que crianças de quatro anos marcaram

com maior altura (*pitch*) e intensidade vocal substantivos que se caracterizavam como informação nova nas sentenças proferidas a partir de um jogo e, em contraposição, marcaram com menor altura (*pitch*) e intensidade substantivos que se caracterizavam como informação dada. Ainda na língua inglesa, outro estudo relata que crianças de quatro anos se utilizaram da frequência fundamental para marcar o foco prosódico contrastivo em situação espontânea (Watson; Arnold; Tanenhaus, 2008).

No Brasil, estudos que tenham contemplado a produção de foco prosódico na fala de crianças falantes do português brasileiro são recentes e, em conjunto, tem apresentado resultados que apontam para as idiosincrasias dos processos de desenvolvimento de fala. Particularmente, pesquisas desenvolvidas no escopo do projeto *Produção e percepção da prosódia em crianças com desenvolvimento fonológico típico e atípico* (FAPESP 2020/10144-3), vêm apresentando resultados importantes acerca da caracterização prosódica na fala de crianças diagnosticadas com TF.

Por exemplo, no trabalho de Soncin, Polli e Berti (2022), no qual foram analisados dados obtidos a partir da aplicação da prova de Acento Frasal<sup>2</sup>, proposta por Preston et al. (2016) como parte de provas que compõem instrumento para avaliação motora da fala, identificou-se que as crianças com TF utilizam pistas fonético acústicas como aumento da duração e da intensidade para marcar foco prosódico contrastivo, mas não utilizam a configuração tonal típica da marcação do foco prosódico no PB.

Por sua vez, Polli (2022) investigou o desempenho prosódico verbal e visual em crianças com TF e crianças em DTL em três diferentes tarefas de produção de fala: acento de palavra, foco prosódico e distinção de fronteira final e não final de enunciado. Dentre outros resultados, o trabalho mostra que esses grupos se diferenciaram entre si, pois o grupo de crianças com TF apresentou desempenho inferior ao grupo de crianças com desenvolvimento típico de linguagem em todas as tarefas. O estudo ainda conclui que o desempenho prosódico verbal e gestual depende de um processo de aquisição, em que o uso de marcadores verbais parece anteceder o uso de marcadores gestuais, porém, a qualidade desse desempenho tem estreita relação com a condição clínica das crianças.

No que diz respeito à comparação da performance de crianças com TF e de crianças com DTL na produção do foco contrastivo, estudo de Santos et al (2023) realizou análise acústica da duração em sentenças com elementos focalizados nas amostras de fala desses grupos de crianças. Com base nos dados obtidos, os resultados permitem concluir que crianças com TF apresentam inconsistências na produção do foco contrastivo que se mostram no aspecto duracional, pois observou-se: (i) tendência ao não aumento de duração em palavras produzidas em contexto de focalização prosódica, diferente do que se observou no grupo de crianças em DTL, em cujas amostras analisadas, o aumento duracional nas palavras focalizadas foi constante; e (ii) mudança da unidade alongada nos casos em que se observou aumento de duração, uma vez que, quando o aumento duracional ocorreu na fala das crianças com TF, ele

---

<sup>2</sup> A prova consistia na repetição de sentenças produzidas com foco prosódico contrastivo em diferentes posições sintáticas, tais como na sentença *João adora jogar bola*. Destaca-se que a produção apresentada como modelo para repetição era a produção do avaliador, ou seja, o avaliador produzia cada sentença e, na sequência, solicitava a repetição do paciente. Por exemplo, o avaliador produzia *JOÃO adora jogar bola* e solicitava a repetição.

não foi realizado na sílaba tônica da palavra alvo do foco, como se espera, de acordo com o padrão de produção típica de crianças e adultos.

Apesar dos resultados apresentados pelos estudos acima reportados, até onde sabemos, não foram realizados trabalhos que tenham comparado como se manifesta o foco prosódico na fala de crianças, tanto em desenvolvimento fonológico típico quanto em desenvolvimento fonológico atípico, e de adultos falantes do português brasileiro, de modo a identificar possíveis diferenças e/ou semelhanças entre esses grupos que permitiriam flagrar marcas dos processos de aquisição prosódica – seja típico ou atípico. Visando preencher essa lacuna, o presente trabalho se desenvolveu.

### 3 MÉTODO

Para atender aos objetivos delineados, o presente trabalho foi desenvolvido à luz de abordagem experimental, com procedimentos comuns à Fonologia de Laboratório (Cohn; Fougeron; Huffman, 2012). Em termos de método, o presente trabalho se caracteriza como um estudo experimental, transversal e quantitativo.

#### 3.1 Aspectos éticos

Este estudo foi aprovado por Comitê de Ética em Pesquisa sob o número 035514/2021 (CAAE: 45522721.6.0000.5406).

#### 3.2 Participantes

Participaram do estudo 30 sujeitos, organizados em três grupos: (1) adultos monolíngues falantes nativos do português brasileiro (PB); (2) crianças em DTL em processo de aquisição monolíngue do PB; (3) crianças com diagnóstico fonoaudiológico de TF em processo de aquisição monolíngue do PB. Cada grupo foi composto por 10 sujeitos.

Os adultos, de faixa etária entre 18 e 24 anos, foram recrutados do curso de graduação em Fonoaudiologia da Universidade Estadual Paulista – câmpus de Marília. As crianças em DTL, de faixa etária entre 5 e 9 anos de idade, foram recrutadas na Escola Municipal de Ensino Fundamental e Educação Infantil Chico Xavier e na Organização Não Governamental Semear, que oferece atividades de formação cultural e esportiva para crianças e adolescentes em contraturno escolar. Ambas se situam na cidade de Marília e são instituições com as quais o projeto de pesquisa ao qual este trabalho se vincula mantém parceria. Por sua vez, as crianças com TF, também de faixa etária entre 5 e 9 anos, foram recrutadas em sessões de avaliação fonoaudiológica, pré-intervenção, durante atendimento realizado em Estágio Supervisionado em Fonoaudiologia com ênfase em Fonologia Clínica oferecido à população em clínica escola da Universidade Estadual Paulista – câmpus de Marília.

Para todos os grupos, foram adotados os seguintes critérios de exclusão: alterações intelectuais e neurológicas; alterações anatomo-morfológicas que comprometessem o processo de produção de fala; alterações otológicas/auditivas. A adoção desses critérios se justifica para que se descartem fatores outros que possam interferir nos resultados da análise dos dados. Especialmente para o grupo de crianças com TF, tais critérios são relevantes para que, em caso de identificação de possíveis diferenças na performance desse grupo, possam ser feitas afirmações que as

relacionem estritamente a aspectos do transtorno de fala, sem associação, portanto, a outras condições clínicas de natureza intelectual, neurológica e/ou fisiológica. Para o grupo de adultos, adotou-se ainda, como critério de inclusão, proveniência de cidades do estado de São Paulo, a fim de garantir identidade quanto à variação linguística dos grupos de crianças. Desse modo, foram utilizados dados de adultos que tivessem em comum a mesma região de procedência das crianças participantes.

### 3.3 Procedimento experimental

Foi desenvolvido um experimento de produção de fala para a produção de sentenças com elementos focalizados e sentenças neutras. As sentenças eram compostas por sujeito, verbo e objeto. Quando produzidas em contexto de focalização prosódica, o foco contrastivo recaía, a cada produção, sobre uma posição sintática<sup>3</sup>. Todos os sujeitos produziram as mesmas sentenças. O quadro abaixo apresenta as sentenças produzidas e indica as palavras focalizadas em maiúsculas.

Quadro 1 - Conjunto de sentenças analisadas.

Sentença	Contexto de produção
As meninas amam vestido vermelho	Sentença neutra
As MENINAS amam vestido vermelho	Foco no sujeito
As meninas AMAM vestido vermelho	Foco no verbo
As meninas amam VESTIDO vermelho	Foco no objeto (posição 1)
As meninas amam vestido VERMELHO	Foco no objeto (posição 2)

Fonte: Elaborado pelos autores.

Para a aplicação do experimento nos grupos de crianças, foi elaborado um jogo de tabuleiro digital que contextualizou a produção das sentenças. Cada casa do tabuleiro apresentava, em formato de áudio, uma sentença associada a uma história narrada anteriormente. Como regra do jogo, para avançar uma casa e pontuar, o participante deveria reproduzir a sentença de maneira idêntica ao que havia ouvido<sup>4</sup>.

<sup>3</sup> No delineamento do experimento, optou-se por variar a posição sintática da palavra alvo de foco contrastivo para que se testasse, em análise específica, se a posição sintática teria efeito no desempenho relativo à produção do foco pelos grupos de crianças. Essa análise, embora tenha sido realizada, foge aos objetivos do presente texto e será alvo de discussão em publicação futura.

<sup>4</sup> Sabemos das limitações associadas a uma tarefa de repetição, especialmente por não se tratar de uma produção contextualizada e, ainda, por ser atravessada pelo aspecto perceptivo do estímulo apresentado para ser reproduzido. No entanto, a adoção de tal procedimento para coleta de dados se justifica por ser esse o procedimento empregado em instrumentos de Avaliação Motora da Fala (Preston et al., 2016) que contempla provas de produção de foco e acento frasal para fins de diagnóstico diferencial, cujos resultados existentes na literatura se restringem a crianças com diagnóstico clínico. Assim, a fim de que se possa comparar a performance de crianças com TF e a performance de crianças e adultos típicos, no presente estudo, decidimos por manter a tarefa de repetição, pois, com os resultados comparativos entre os grupos aqui contemplados teríamos melhores condições de discutir nossos achados com aqueles já reportados na literatura, haja vista a similaridade das tarefas. Contudo, salientamos que novos dados estão sendo coletados pela equipe de pesquisa a partir de procedimento experimental baseado em trocas comunicativas em contextos específicos que elicitam a produção de foco contrastivo de forma mais espontânea (ver por exemplo, Lima; Guida; Soncin, 2023).

Venceria o jogo o participante que somasse os pontos e cruzasse a linha de chegada. Para o grupo de adultos, por sua vez, o experimento foi aplicado sem o recurso do jogo de tabuleiro para apresentar a tarefa de modo mais adequado a esse público, sem o uso de narrativas e personagens infantis. Para esse grupo, as mesmas sentenças usadas no jogo foram apresentadas aos participantes adultos por meio de arquivos de áudio inseridos em apresentação de slides, a partir dos quais solicitou-se a repetição<sup>5</sup>. Nos slides, nenhuma informação gráfica das sentenças apresentadas auditivamente para repetição foi inserida.

As sentenças apresentadas como modelo foram gravadas previamente em cabine de isolamento acústico por um falante adulta nativa do PB, de sexo feminino, idade de 21 anos e proveniente da cidade de Marília, interior de São Paulo. A produção de cada participante durante os experimentos foi gravada com software Samsung Recorder Plus e microfone Sony, modelo ECM-CS3, e submetida à análise acústica.

### 3.3 Análise acústica

As sentenças produzidas foram analisadas com o software *Praat* (Boersma; Weenink, 2022). No total, 150 sentenças foram analisadas, sendo 50 de cada grupo (5 sentenças x 10 participantes x 3 grupos). Três parâmetros foram considerados na análise acústica para a caracterização da marcação de foco prosódico contrastivo: duração, intensidade e entoação. Para análise da duração, mensurou-se a duração relativa da sílaba tônica da palavra alvo de foco prosódico e a duração relativa da sílaba tônica da mesma palavra quando produzida em sentença neutra (por exemplo, considerando o Quadro 1, mensurou-se a duração relativa da sílaba tônica de *meninas* na sentença com foco no sujeito e na sentença neutra). Assim, comparou-se a duração do mesmo elemento em dois contextos prosódicos distintos – com foco prosódico e sem foco prosódico –, a fim de verificarmos se haveria diferença na duração entre os contextos. Para análise da intensidade, mensurou-se, em decibéis, o valor de intensidade máxima na sílaba tônica da palavra alvo de foco prosódico e a intensidade da mesma sílaba fora de contexto de focalização para fins de comparação. Para análise da entoação, considerou-se tanto o valor de F0 quanto a caracterização tonal das palavras focalizadas nas sentenças. Na mensuração do valor de F0, extraiu-se em Hertz o valor do pico de F0 na sílaba tônica da palavra focalizada bem como na sílaba tônica da mesma palavra fora de contexto de focalização, de modo similar ao realizado para a intensidade. Por sua vez, na caracterização tonal, identificou-se o tipo de configuração tonal associada ao elemento focalizado realizada pelos participantes a fim

<sup>5</sup> Destaca-se que a mudança no procedimento realizado com o grupo de adultos em relação aos grupos de crianças, ou seja, a não utilização do jogo de tabuleiro para o grupo de adultos, justifica-se pela adequação à faixa etária das crianças, já que o jogo faz uso de personagens infantis e foi desenvolvido para atender às demandas de atenção das crianças durante um procedimento experimental. Embora o procedimento, em termos de suporte, seja distinto, salientamos que a tarefa solicitada a adultos e crianças foi a mesma: a reprodução de áudios apresentados na ferramenta em uso, seja ela o jogo de tabuleiro digital ou o material de apresentação em *Power Point*. Como já mencionado anteriormente, sabemos das limitações associadas a uma tarefa de repetição, especialmente por não se tratar de uma produção contextualizada e, ainda, por ser atravessada pelo aspecto perceptivo do estímulo apresentado para ser reproduzido. Para trabalhos do grupo de pesquisa que estão em curso de desenvolvimento, no entanto, estão sendo aplicados novos procedimentos experimentais para coleta de dados de produção de foco contrastivo e informacional, baseados em tarefas que consistem em trocas comunicativas em contextos específicos e – portanto – não baseadas em tarefas de repetição. Os resultados poderão ser acessados em publicações futuras.

de verificar se corresponderia ao padrão característico descrito para foco contrastivo no PB (Frota et al., 2015; Yano; Fernandes-Svartman, 2020).

### 3.4 Análise estatística

Os dados das três pistas fonéticas receberam tratamento estatístico descritivo e inferencial. Aplicou-se o teste T Pareado para a análise de cada parâmetro nas amostras dos diferentes grupos, considerando os valores dos parâmetros como variável dependente e o contexto de produção como variável independente. Estabeleceu-se o nível de significância  $\alpha = 5\%$  e intervalo de confiança IC = 95% ( $p\text{-value} < 0,05$ ).

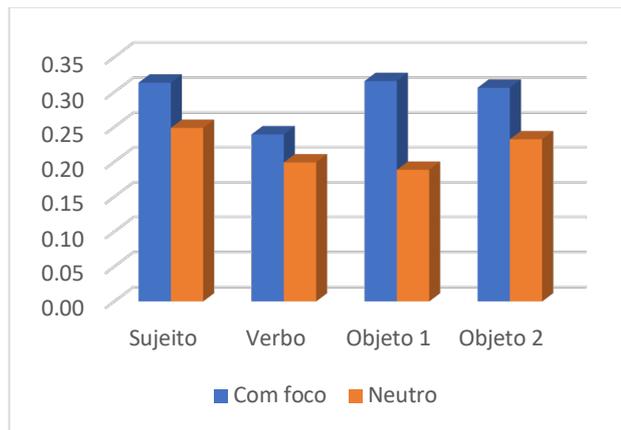
## 4 RESULTADOS E DISCUSSÃO

Para o grupo de adultos, foram obtidos os seguintes resultados: (i) no que diz respeito à duração, o teste indicou maior duração da sílaba tônica das palavras em contexto de focalização prosódica em todas as posições sintáticas onde recaiu o foco prosódico contrastivo ( $p < 0,05$ ) (Gráfico 1)<sup>6</sup>; (ii) quanto à intensidade, o teste indicou maior pico de intensidade na sílaba tônica da palavra onde recaiu o foco prosódico contrastivo para duas posições sintáticas, sendo elas verbo e segunda posição do objeto ( $p < 0,05$ ) (Gráfico 2)<sup>7</sup>; (iii) no que diz respeito à entoação, o teste indicou maior magnitude de F0 na sílaba tônica da palavra quando em contexto de focalização prosódica contrastiva para todas as posições sintáticas ( $p < 0,05$ ) (Gráfico 3)<sup>8</sup>, e observou-se ainda que os acentos nucleares associados aos elementos focalizados produzidos corresponderam em 100% ao padrão entoacional de foco contrastivo do PB, sendo eles: L\*+H seguido de tom de fronteira H%, quando a palavra focalizada ocorreu em posição não final da sentença; e H+L\* seguido de tom de fronteira L%, quando a palavra focalizada ocorreu em posição final. As configurações tonais identificadas no grupo de adultos corroboram, assim, a caracterização entoacional de sentenças com elementos focalizados contrastivamente apresentada por Yano e Fernandes-Svartman (2020) e Frota et al. (2015).

<sup>6</sup> Estatística t: Foco no sujeito,  $t(9) = 5,110$ ,  $p = 0,001$ ; Foco no verbo,  $t(9) = 2,549$ ,  $p = 0,031$ ; Foco na posição 1 de objeto,  $t(9) = 10,483$ ,  $p = 0,000$ ; Foco na posição 2 de objeto  $t(9) = 3,633$ ,  $p = 0,005$ .

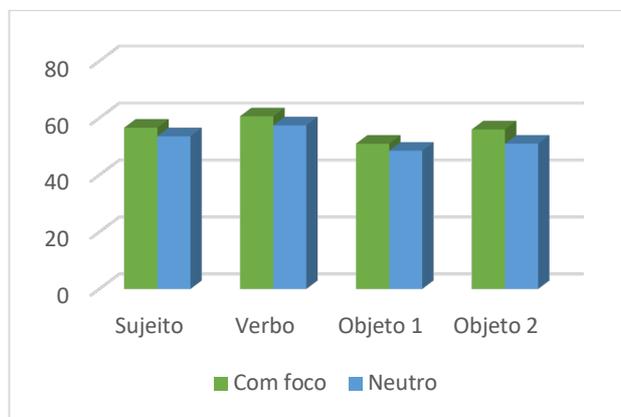
<sup>7</sup> Estatística t: Foco no sujeito,  $t(9) = 1,598$ ,  $p = 0,144$ ; Foco no verbo,  $t(9) = 3,309$ ,  $p = 0,009$ ; Foco na posição 1 de objeto,  $t(9) = 0,886$ ,  $p = 0,398$ ; Foco na posição 2 de objeto  $t(9) = 5,783$ ,  $p = 0,000$ .

<sup>8</sup> Estatística t: Foco no sujeito,  $t(9) = 4,884$ ,  $p = 0,001$ ; Foco no verbo,  $t(9) = 6,435$ ,  $p = 0,000$ ; Foco na posição 1 de objeto,  $t(9) = 5,077$ ,  $p = 0,001$ ; Foco na posição 2 de objeto  $t(9) = 3,885$ ,  $p = 0,004$ .



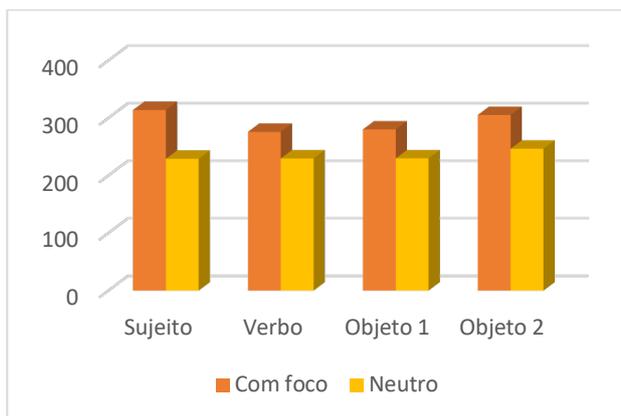
Fonte: Elaborado pelos autores.

Gráfico 1 - Média de duração por contexto no grupo de adultos.



Fonte: Elaborado pelos autores.

Gráfico 2 - Média do pico de intensidade por contexto no grupo de adultos.

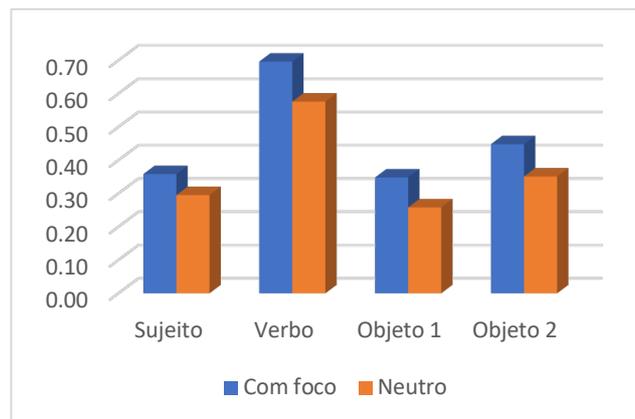


Fonte: Elaborado pelos autores.

Gráfico 3 - Média da magnitude de F0 por contexto no grupo dos adultos.

Para o grupo de crianças em DTL, os testes indicaram os seguintes resultados: em relação à duração, houve maior duração da sílaba tônica das palavras em contexto de focalização contrastiva em três das quatro posições sintáticas consideradas, nomeadamente, no verbo e nas posições 1 e 2 de objeto ( $p < 0,05$ ), exceto na posição

de sujeito (Gráfico 4)<sup>9</sup>. Em relação à intensidade, houve maior pico na sílaba tônica da palavra focalizada, apenas quando o foco prosódico contrastivo recaiu sobre a segunda posição do objeto ( $p < 0,05$ ) (Gráfico 5)<sup>10</sup>. Quanto à entoação, ocorreu maior magnitude de F0 nas sílabas tônicas das palavras, quando em contexto de focalização prosódica contrastiva, para as posições sintáticas de sujeito e verbo ( $p < 0,05$ ), não tendo sido observada diferença estatisticamente relevante de F0 entre os contextos comparados (neutro e com foco contrastivo) nas posições 1 e 2 de objeto, (Gráfico 6)<sup>11</sup>. Observou-se ainda que o percentual de configuração tonal associada aos elementos focalizados realizado de acordo com o padrão entoacional de foco contrastivo do PB foi equivalente a 75%. Assim como para o grupo de adultos, os acentos nucleares associados às palavras focalizadas produzidos em conformidade com o padrão entoacional de foco contrastivo foram L\*+H seguido de tom de fronteira H%, quando o foco foi realizado em posição não final de enunciado, e, também, H+L\* seguido de L%, quando o foco foi produzido em posição final. Por sua vez, em relação aos 25% que foram considerados fora do padrão entoacional do PB, identificou-se, predominantemente, H\* associado aos elementos focalizados produzidos em posição não final de enunciado e L\* seguido de L% associado a elemento focalizado em posição final de enunciado. Essas configurações tonais foram consideradas fora do padrão entoacional do PB por não terem sido reportadas por descrições entoacionais anteriores como representativas do foco contrastivo nessa variedade de português.



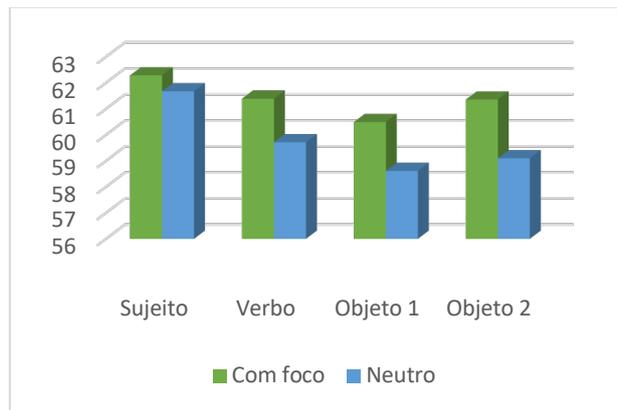
Fonte: Elaborado pelos autores.

Gráfico 4 - Média de duração por contexto no grupo de crianças com DTL.

<sup>9</sup> Estatística t: Foco no sujeito,  $t(9) = 1,909$ ,  $p = 0,089$ ; Foco no verbo,  $t(9) = 3,547$ ,  $p = 0,006$ ; Foco na posição 1 de objeto,  $t(9) = 3,468$ ,  $p = 0,007$ ; Foco na posição 2 de objeto  $t(9) = 5,403$ ,  $p = 0,000$ .

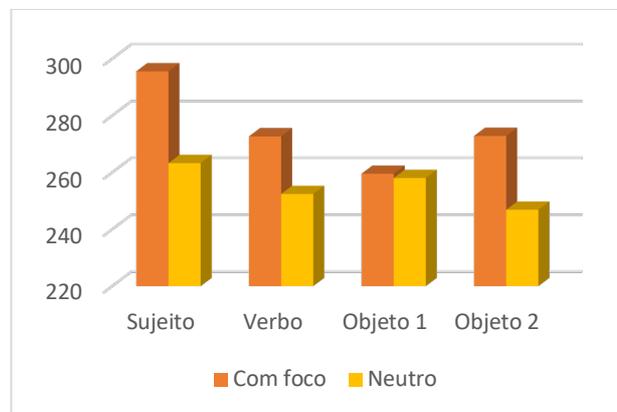
<sup>10</sup> Estatística t: Foco no sujeito,  $t(9) = 0,651$ ,  $p = 0,531$ ; Foco no verbo,  $t(9) = 1,652$ ,  $p = 0,133$ ; Foco na posição 1 de objeto,  $t(9) = 2,167$ ,  $p = 0,058$ ; Foco na posição 2 de objeto  $t(9) = 2,972$ ,  $p = 0,016$ .

<sup>11</sup> Estatística t: Foco no sujeito,  $t(9) = 3,596$ ,  $p = 0,006$ ; Foco no verbo,  $t(9) = 2,333$ ,  $p = 0,045$ ; Foco na posição 1 de objeto,  $t(9) = 0,078$ ,  $p = 0,939$ ; Foco na posição 2 de objeto  $t(9) = 2,071$ ,  $p = 0,068$ .



Fonte: Elaborado pelos autores.

Gráfico 5 - Média do pico de intensidade por contexto no grupo de crianças em DTL.



Fonte: Elaborado pelos autores.

Gráfico 6 - Média da magnitude de F0 por contexto no grupo de crianças em DTL.

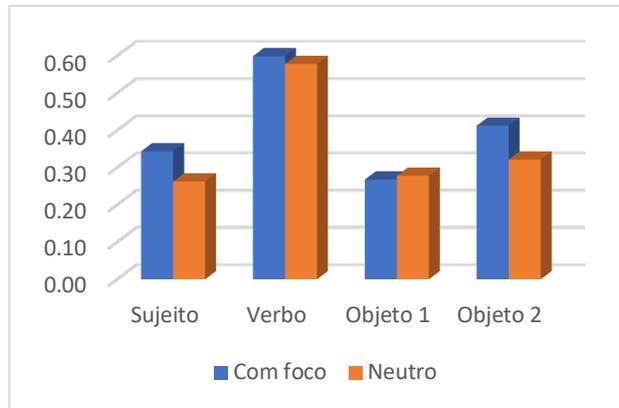
Para o grupo de crianças com TF, foram obtidos os seguintes resultados: (i) houve maior duração estatisticamente relevante na sílaba tônica da palavra focalizada apenas na posição 2 de objeto ( $p < 0,05$ ), não tendo sido identificada diferença de duração na sílaba tônica da palavra focalizada nas demais posições sintáticas (Gráfico 7)<sup>12</sup>; (ii) identificou-se maior intensidade e maior magnitude de F0 na sílaba tônica da palavra focalizada de forma estatisticamente significativa, apenas quando o foco contrastivo recaiu no verbo ( $p < 0,05$ ) (Gráficos 8<sup>13</sup> e 9<sup>14</sup>); (iii) na caracterização tonal, o percentual de configuração tonal realizada de acordo com o padrão entoacional de foco contrastivo do PB correspondeu a 70%, sendo L\*+H seguido de tom de fronteira H% e H+L\* L% as configurações tonais identificadas em palavras focalizadas em posição não final e final, respectivamente. Em contrapartida, compuseram os 30% de configurações tonais consideradas fora do padrão entoacional do PB para foco

<sup>12</sup> Estatística t: Foco no sujeito,  $t(8) = 2,244$ ,  $p = 0,055$ ; Foco no verbo,  $t(8) = 0,487$ ,  $p = 0,640$ ; Foco na posição 1 de objeto,  $t(8) = -0,476$ ,  $p = 0,647$ ; Foco na posição 2 de objeto  $t(8) = 3,159$ ,  $p = 0,013$ .

<sup>13</sup> Estatística t: Foco no sujeito,  $t(8) = 0,224$ ,  $p = 0,829$ ; Foco no verbo,  $t(8) = 2,792$ ,  $p = 0,023$ ; Foco na posição 1 de objeto,  $t(8) = 0,094$ ,  $p = 0,928$ ; Foco na posição 2 de objeto  $t(8) = 0,736$ ,  $p = 0,483$ .

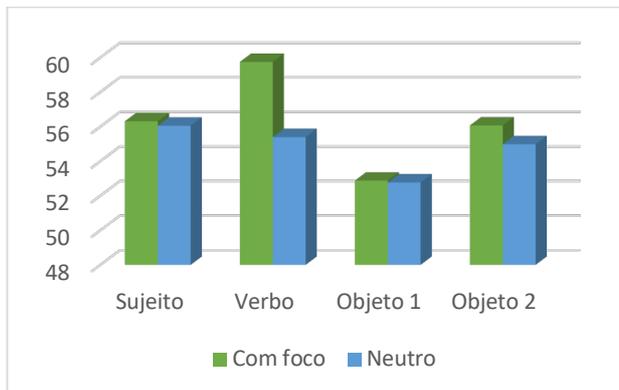
<sup>14</sup> Estatística t: Foco no sujeito,  $t(8) = 0,505$ ,  $p = 0,627$ ; Foco no verbo,  $t(8) = 3,074$ ,  $p = 0,015$ ; Foco na posição 1 de objeto,  $t(8) = -0,060$ ,  $p = 0,954$ ; Foco na posição 2 de objeto  $t(8) = 1,623$ ,  $p = 0,143$ .

contrastivo os acentos L\* e H\*+L, ambos realizados em posição não final. Não foram identificados tons de fronteira nesses casos.



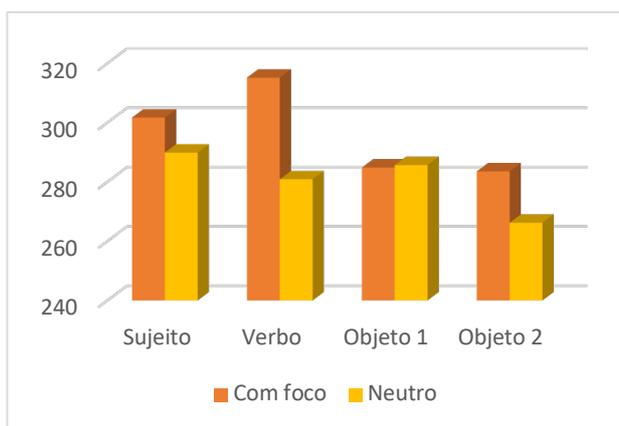
Fonte: Elaborado pelos autores.

Gráfico 7 - Média de duração por contexto no grupo de crianças com TF.



Fonte: Elaborado pelo autor.

Gráfico 8- Média do pico de intensidade por contexto no grupo de crianças com TF.



Fonte: Elaborado pelo autor.

Gráfico 9 - Média da magnitude de F0 por contexto no grupo de crianças com TF.

Os resultados permitem observar que há diferenças no modo como adultos, crianças em DTL e crianças com TF produzem o foco prosódico contrastivo, como

hipotetizado. Tais diferenças se mostram nos três parâmetros acústicos considerados: duração, intensidade e entoação.

No grupo de adultos, considerado controle em relação aos demais grupos, o foco contrastivo foi marcado sistematicamente por aumento de duração, aumento da magnitude de F0 e pela produção de configuração tonal típica de foco prosódico no PB em todas as posições sintáticas, enquanto o aumento de intensidade não foi regular. Tais resultados corroboram achados anteriores já apresentados pela literatura sobre o papel dos parâmetros na marcação do foco, os quais indicam o aumento de intensidade como o parâmetro menos robusto dentre os três e a F0 como o mais robusto entre eles, seguido da duração (Moraes, 2009).

No grupo de crianças em DTL, a duração foi o parâmetro mais consistentemente usado, tendo sido observado aumento nas diferentes posições sintáticas, exceto no sujeito. Diferentemente do grupo de adultos, duração e entoação não foram usadas de forma combinada sistematicamente, mas observou-se relação de complementaridade no uso desses dois parâmetros nas diferentes posições sintáticas, uma vez que, no sujeito, única posição não marcada por aumento de duração por esse grupo, houve maior magnitude de F0, além de configuração tonal prototípica de foco prosódico. Dessa observação, constata-se que esse grupo de sujeitos produziu o foco prosódico em todas as posições sintáticas, marcando-o, ao menos, por meio de alguma pista acústica.

Por um lado, esses resultados mostram que crianças menores de nove anos falantes do PB produzem foco prosódico contrastivo, ainda que sua produção não esteja conforme o padrão adulto – o que já seria esperado. Desse modo, os resultados apresentados questionam os trabalhos que afirmam que a aquisição da produção do foco contrastivo se daria após os nove anos de idade (cf. para o inglês britânico, Wells; Peppé; Goulandris, 2004; cf. para o inglês da Nova Zelândia, Kalathottukaren; Purdy, 2017; cf. para o português europeu, Filipe et al., 2017). Por outro lado, esses resultados, para o grupo de crianças em DTL, é consonante com os trabalhos de Watson, Arnold e Tanenhaus (2008) e Wonnacott e Watson (2008), cujos resultados mostrados para a língua inglesa indicam que crianças em DTL com menos de nove anos de idade (esses estudos contemplam crianças com quatro anos) diferenciam elementos focalizados e elementos neutros; diferenças essas que se podem observar por meio de detalhamento acústico.

Em contrapartida, o grupo de crianças com TF apresentou indícios de marcação de foco prosódico contrastivo apenas em duas posições sintáticas: na segunda posição de objeto, por meio do aumento da duração, e no verbo, por meio de maior magnitude de F0 e de intensidade. Observou-se ainda que esse grupo de sujeitos apresentou maior variabilidade de configurações tonais fora do padrão do PB associada ao elemento focalizado. Desse modo, o grupo com TF apresentou, em alguma medida, inconsistências na marcação de foco prosódico contrastivo no uso dos três parâmetros e defasagem em relação ao grupo de crianças em DTL. Trata-se de resultado relevante, na medida em que corrobora o que vem sendo apontado por outros estudos, como de Polli (2022), Soncin, Polli e Berti (2022) e Santos et al. (2023): maior atenção deve ser dada às características prosódicas das crianças diagnosticadas com TF, uma vez que possíveis inconsistências prosódicas apresentadas na fala dessas crianças podem indicar dificuldades na organização de um sistema prosódico também em aquisição durante o processo de aquisição fonológica.

## 6 CONCLUSÃO

No presente trabalho, as diferenças no uso de parâmetros acústicos nos três grupos contemplados são vistas como indícios de processo(s) de aquisição do foco prosódico contrastivo. Destaca-se, porém, que esses processos são afetados pelas condições desenvolvimental e clínica dos sujeitos. Diferente dos adultos, que marcaram o foco contrastivo sistematicamente por meio de aumento de duração, aumento da magnitude de F0 e pela configuração tonal típica de foco prosódico no PB, crianças em DTL produziram foco prosódico privilegiadamente por meio de aumento da duração. Destaca-se ainda que, nos dados desse grupo, quando o aumento duracional não ocorreu na palavra focalizada, observou-se maior magnitude de F0 e configuração tonal típica de foco, mostrando que, na fala de crianças em DTL, ainda que a produção do foco não seja prototípica, conforme observado nos dados dos adultos, a mudança no comportamento dos parâmetros registra indícios da produção da focalização. Por sua vez, crianças com TF oscilaram no uso dos parâmetros e não apresentaram uma regularidade que caracterizaria a marcação de foco prosódico. Nas amostras de fala dessas crianças, diferentemente dos dados de crianças em DTL, não se observou aumento duracional na palavra focalizada de forma recorrente e significativa. Assim, o estudo sugere que: por um lado, crianças menores de nove anos em DTL produzem foco prosódico contrastivo de forma gradiente em relação aos adultos; e, por outro, crianças com TF, de mesma idade, podem apresentar defasagem prosódica na produção de fala. Como implicação prática desses resultados para o contexto clínico, constata-se a necessidade de desenvolvimento de instrumentos de avaliação de fala que contemplem aspectos prosódicos, de modo a identificar essas e outras possíveis alterações nesse plano fonético-fonológico, dado o impacto que a prosódia tem para habilidades comunicativas.

## REFERÊNCIAS

- American Psychiatric Association. DSM-5: Manual diagnóstico e estatístico de transtornos mentais. Porto Alegre: Artmed Editora; 2014.
- American Speech-Language-Hearing Association. Speech sound disorders - Articulation and phonology [internet]. 2023.
- Arvaniti A, Fletcher J. The autosegmental-metrical theory of intonational phonology. In: Gussenhoven C, Chen A, editores. The Oxford handbook of language prosody. Oxford: Oxford University Press; 2020. p. 78-95.
- Barbosa P. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. *Revista Estudos da Linguagem*. 2012;20(3):11-27.
- Barbosa P, Madureira S. Manual de fonética acústica experimental: aplicações a dados do português. São Paulo: Cortez; 2015.
- Beckman M, Pierrehumbert J. Intonational structure in Japanese and English. *Phonology Yearbook*. 1986;3:255-309.
- Berti LC, Pagliuso A, Lacava F. Instrumento de avaliação de fala para análise acústica (IAFAC) baseado em critérios linguísticos. *Revista da Sociedade Brasileira de Fonoaudiologia*. 2009;14(3):305-314.
- Berti LC, et al. Relationship between speech production and perception in children with speech sound disorders. *Journal of Portuguese Linguistics*. 2020;19(13):1-13.
- Berti LC, et al. Differentiation on linguistic manifestations in children with delayed phonological development; em preparação.

- Boersma P, Weenink D. Praat: doing phonetics by computer [programa de computador]. Versão 6. 2.1.4. Amsterdam: Universiteit van Amsterdam; 2022.
- Carnaval M, Moraes JA, Rilliard A. Focus types in Brazilian Portuguese: multimodal production and perception. *DELTA*. 2022;38(3):202258944.
- Ceron MI, et al. Instrumento de avaliação fonológica (INFONO): estudo piloto. *Codas*. 2020;4(2):1-13.
- Chafe WL. Givenness, contrastiveness, definiteness, subjects and topics. In: Li CN, editor. *Subject and topic*. New York: Academic Press; 1976. p. 27-55.
- Cohn CA, Fougeron C, Huffman, MK. *The Oxford handbook of laboratory phonology*. Oxford: Oxford University Press; 2012.
- D'Imperio M, et al. Intonational phrasing in Romance: the role of syntactic and prosodic structure. In: Frota S, Vigário M, Freitas MJ, editores. *Prosodies*. Berlin: Mouton de Gruyter; 2005. p. 59-97.
- De Ruiter JP. *Questions: formal, functional and interactional perspectives*. Cambridge: Cambridge University Press; 2012.
- Diehl JJ, Paul R. The assessment and treatment of prosodic disorders and neurological theories of prosody. *International Journal of Speech Language Pathology*. 2009;11(4):287-292.
- Dodd B. *Differential diagnosis and treatment of children with speech disorder*. 2.<sup>a</sup> ed. London: Whurr; 2005.
- Fernandes FR. *Ordem, focalização e preenchimento em português: sintaxe e prosódia [tese]*. Campinas: Instituto de Estudos da Linguagem, Universidade Estadual de Campinas; 2007.
- Filipe M, et al. Prosodic development in European Portuguese from childhood to adulthood. *Applied Psycholinguist*. 2017;38(5):1045-1070.
- Frota S, et al. Intonational variation in Portuguese: European and Brazilian varieties. In: Frota S, Prieto P, organizadores. *Intonation in Romance*. Oxford: Oxford University Press; 2015. p. 235-283.
- Geluykens R, Swerts M. Prosodic cues to discourse boundaries in experimental dialogues. *Speech Communication*. 1994;15(1-2):69-77.
- Gonçalves CAV. *Focalização no português do Brasil [tese]*. Rio de Janeiro: Faculdade de Letras, Universidade Federal do Rio de Janeiro; 1997.
- Gonçalves GF, Keske-Soares M, Checalin MA. Estudo do papel do contexto linguístico no tratamento do desvio fonológico. *Revista da Sociedade Brasileira de Fonoaudiologia*. 2010;15(1):96-102.
- Gussenhoven C. Types of focus in English. In: Lee C, Gordon M, Buring D, editores. *Topic and focus: cross-linguistic perspectives on meaning and intonation*. Dordrecht: Springer; 2006. p. 83-100.
- Gussenhoven C. Sentential prominence in English. In: Ostendorp M et al., editores. *The Blackwell companion to phonology*. Nova Jersey: Wiley-Blackwell; 2011. p. 2780-2806.
- Hawthorne K, Fischer S. Speech-language pathologists and prosody: clinical practices and barriers. *Journal of Communication Disorders*. 2020;87:106024.
- Hearnshaw S, Baker E, Munro, N. The speech perception skills of children with and without speech sound disorder. *Journal of Communication Disorders*. 2018;71:61-71.
- Kalathottukaren RT, Purdy SC. Prosody perception in typically developing school-aged children. *Journal Phonetics and Audiology*. 2017;3(1):1-9.
- Krifka M. Basic notions of information structures. *Acta Linguistica Hungarica*. 2008;55(3):243-276.
- Kugler F, Calhoun S. Prosodic encoding of information structure: a typological perspective. In: Gussenhoven C, Chen A, editores. *The Oxford handbook of prosody language*. New York: Oxford University Press; 2020. p. 454-467.

- Ladd DR, Cutler A. Introduction. Models and measurements in the study of prosody. In: Cutler A, Ladd DR, editores. *Prosody: models and measurements*. Berlin, Heidelberg: Springer Berlin Heidelberg; 1983. p. 1-10.
- Ladd DR. *Intonational phonology*. Cambridge: Cambridge University Press; 2008.
- Lamprecht RR. *Aquisição fonológica do português: perfil de desenvolvimento e perfil para terapia*. Porto Alegre: Artmed Editora; 2004.
- Lima FLCN, Guida CLS, Soncin G. Julgamento perceptivo-auditivo de habilidades prosódicas na fala de crianças com distúrbios dos sons da fala. *Anais do 31º Congresso Brasileiro de Fonoaudiologia e 12º Congresso Internacional de Fonoaudiologia*; 03-06 out. 2023; Rio de Janeiro, RJ, Brasil. Rio de Janeiro: Sociedade Brasileira de Fonoaudiologia; 2023. p. 1179-1182.
- Moraes JA. Three types of prosodic focus in Brazilian Portuguese: form and meaning; 2009; Workshop on prosody and meaning abstracts; 2009 September 17-18; Barcelona, Espanha. Barcelona: Publisher Unknown; 2009, p. 59-60.
- Nespor M, Vogel I. *Prosodic phonology*. Dordrecht: Foris; 1986.
- Nespor M, Vogel I. *Prosodic phonology: with a new foreword*. Berlin: Mouton de Gruyter; 2007.
- Patah LK, Takiuchi N. Prevalência das alterações fonológicas e uso dos processos fonológicos em escolares aos 7 anos. *Revista CEFAC*. 2008;10(2):158-167.
- Peppé S. Why is prosody in speech-language pathology so difficult? *International Journal of Speech-Language Pathology*. 2009;11(4):258-271.
- Peppé S, McCann J. Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*. 2003;17(4-5):345-354.
- Pierrehumbert J. *The phonology and phonetics of English intonation [tese]*. Massachusetts: Department of Linguistics and Philosophy, Massachusetts Institute of Technology; 1980.
- Polli L. *Desempenho prosódico verbal e gestual em crianças com e sem diagnóstico de distúrbios dos sons da Fala [dissertação]*. Marília: Faculdade de Filosofia e Ciências, Universidade Estadual Paulista; 2022.
- Preston JL, et al. Limited acquisition and generalisation of rhotics with ultrasound visual feedback in childhood apraxia. *Clinical Linguistics & Phonetics*. 2016;30(3-5):363-381.
- Santos KA, et al. Focalização prosódica na fala de crianças com desenvolvimento fonológico típico e atípico: análise duracional. *Veredas - Revista de Estudos Linguísticos*. 2023;27(1):e40928.
- Shriberg LD, et al. Reference data for children with idiopathic speech delay with and without speech motor delay (SMD). Madison: University of Wisconsin-Madison; 2019. Technical Report No. 26, Phonology Project.
- Shriberg LD, Aram D, Kwaitkowski J. Developmental apraxia of speech: III. A subtype marked by inappropriate stress. *Journal of Speech, Language, and Hearing Research*. 1997;40(2):313-337.
- Soncin G, Polli L, Berti LC. Use of secondary cues in prosodic focus marking in speech of children with phonological disorder. *DELTA*. 2022;38(3):202258881.
- Terken J, Hermes D. The perception of prosodic prominence. In: Horne M, editor. *Prosody: theory and experiment: studies presented to Gösta Bruce*. Dordrecht: Springer Netherlands; 2000. p. 89-127.
- Watson DG, Arnold JE, Tanenhaus MK. Tic Tac TOE: effects of predictability and importance on acoustic prominence in language production. *Cognition*. 2008;106(3):1548-1557.
- Wells B, Peppé S, Goulandris N. Intonation development from five to thirteen. *Journal of Child Language*. 2004;31(4):749-778.
- Wertzner HF. Distúrbio fonológico. In: Limongi SCO, editor. *Linguagem: desenvolvimento normal, alterações e distúrbios: fonoaudiologia informação para a formação*. Rio de Janeiro: Guanabara Koogan; 2003. p.1-18.
- Wonnacott E, Watson DG. Acoustic emphasis in four-year-olds. *Cognition*. 2008;107(3):1093-1101.

Yano CT, Fernandes-Svartman FR. Um estudo preliminar sobre a prosódia de construções com tópico e foco no português paulista. *Entrepalavras*. 2020;10(1):256-282.

Yavas M. Padrões na aquisição da fonologia do português. *Letras de Hoje*. 1988;23(4):7-30.



# Os impactos vocais da doença de Parkinson em idosos

## *The vocal impacts of Parkinson's disease in the elderly*

Lucas Manca Dal'Ava\*

*Universidade Estadual de Campinas, Campinas, SP, Brasil*

Plínio Almeida Barbosa\*\*

*Universidade Estadual de Campinas, Campinas, SP, Brasil*

**Resumo:** Este estudo teve como objetivo analisar as diferenças nos parâmetros acústicos de diferentes dimensões fonéticas em discursos espontâneos e leitura em voz alta entre idosos, com foco específico em avaliar o impacto da doença de Parkinson (DP) e do envelhecimento saudável nos biomarcadores de fala de idosos, incluindo parâmetros acústico-prosódicos. Analisamos gravações de 32 falantes idosos brasileiros nativos, considerando três parâmetros acústico-prosódicos: frequência fundamental, duração e intensidade. Nossos resultados sugerem que os fatores *estilo de fala* e *estágios de gravidade da DP* estão correlacionados com a taxa de articulação, inclinação de LTAS em frequências médias, *shimmer*, mediana de F0, mínimo de F0, desvio padrão da primeira derivada positiva de F0 e coeficiente de variação da intensidade.

**Palavras-chave:** Biomarcadores de fala. Doença de Parkinson. Análise acústica. Envelhecimento vocal.

**Abstract:** This study aimed to assess differences in acoustic parameters from different phonetic dimensions in spontaneous and reading aloud speeches among the elderly, with a specific focus on evaluating the impact of Parkinson's disease and healthy aging on speech biomarkers of elderly individuals, including acoustic-prosodic parameters. We analyzed recordings from 32 native Brazilian speakers considering three prosodic-acoustic parameters: fundamental frequency, duration, and intensity. Our results suggests that the factors *speaking style*, and *PD severity stages* are correlated to the articulation rate, medial-frequency long-term average spectrum (LTAS) slope, shimmer, F0 median, F0 minimum, standard deviation of the first positive derivative of F0, and intensity coefficient of variation.

**Keywords:** Speech biomarkers. Parkinson's disease. Acoustic analysis. Vocal aging.

## 1 INTRODUÇÃO

Neste estudo, investigamos os acometimentos fônicos, especificamente prosódicos, da doença de Parkinson (DP) na população idosa brasileira do estado de São Paulo. Essa doença afeta de 1% a 2% da população idosa mundial, com idade acima de 65 anos (Zardeto-Sabec et al., 2018) e pode ser considerada a segunda doença

---

\* Doutorando do Programa de Pós-Graduação em Linguística, Departamento de Linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, SP, Brasil; lucasmdalava@gmail.com

\*\* Professor Titular do Departamento de Linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, SP, Brasil; pbarbosa@unicamp.br

neurodegenerativa mais comum em idosos no mundo (Fernandes; Andrade Filho, 2018). A prevalência na população brasileira é estimada em 3,3% (Barbosa et al., 2006).

Por ser uma doença que afeta principalmente os idosos, as primeiras alterações fônicas podem ser confundidas com as características do envelhecimento natural da voz, presbifonia. Tanto para os idosos saudáveis quanto para os diagnosticados com DP, espera-se a presença de sintomas como tremor e rigidez muscular (Ikuta et al., 2012). Espera-se que as alterações fônicas se agravem gradualmente com o avanço da idade e estágios de severidade da DP (Hoehn; Yahr, 1967).

Para estudar os biomarcadores<sup>1</sup> de parâmetros prosódico-acústicos de idosos com DP (Appakaya; Sankar, 2018), é possível analisar as variações de frequência fundamental (F0) para auxiliar no diagnóstico diferencial e durante o acompanhamento clínico (Dao et al., 2022). Dentre os parâmetros fônicos, a prosódia é o aspecto mais relevante, pois traz informações dos aspectos biológicos, psicológicos e sociais do falante, nos levando a uma descrição abrangente da fala (Lirani-Silva; Mourão; Gobbi 2015). A prosódia é o conjunto de características suprasegmentais da fala, abrangendo entonação, ritmo, duração, intensidade e pausas, além de dimensões linguísticas, cognitivas e sociais. Linguisticamente, a prosódia organiza a fala em constituintes e permite chamar a atenção para unidades linguísticas por sua função de proeminência, bem como distinguir tipos de frase. Cognitivamente, envolve a percepção e interpretação pelo ouvinte, como compreender intenções comunicativas. Socialmente, a prosódia transmite identidade e regula interações sociais (Ladd, 2019). Acrescenta-se que a seleção dos parâmetros usados como biomarcadores carece de sistematização e existem poucos trabalhos em língua portuguesa abordando esse tema.

Neste estudo, foram considerados três parâmetros prosódico-acústicos para analisar as alterações fônicas dos idosos: F0 (correlato acústico da vibração das pregas vocais e sua variabilidade é responsável pela entonação), duração (correlato físico correspondente ao tempo de articulação de uma unidade linguística) e intensidade (correlato psico-acústico da amplitude).

As alterações fônicas em pessoas idosas em processo de envelhecimento saudável estão principalmente relacionadas ao próprio processo e são predominantemente graduais, envolvendo a perda de elasticidade nas pregas vocais, redução na frequência fundamental e outros fatores associados à idade. As alterações fônicas na DP são diretamente ligadas a distúrbios neurológicos específicos, como bradicinesia, rigidez muscular e incoordenação pneumofonoarticulatória. A bradicinesia na DP resulta em diminuição da F0 e da intensidade da voz, acompanhada de alterações na taxa de articulação<sup>2</sup> devido à lentidão na execução de movimentos

---

<sup>1</sup> Os biomarcadores de parâmetros prosódico-acústicos ou da fala podem ser analisados por meio de metodologia específica e comparados com valores de referência de uma determinada população, como funcionam os exames clínicos laboratoriais. Esses biomarcadores são medidas ou características específicas da fala que podem ser usadas para identificar, monitorar ou avaliar condições médicas, distúrbios neurológicos, distúrbios da comunicação ou outros estados de saúde. Esses biomarcadores são baseados em análises quantitativas e qualitativas da fala, que podem incluir elementos como padrões de entonação, taxa de elocução, articulação, volume, ritmo e outros aspectos acústicos e linguísticos da comunicação verbal.

<sup>2</sup> A taxa de articulação mede a taxa de produção dos sons durante a fala sem considerar as pausas silenciosas, possibilitando, assim, revelar mais claramente aspectos como a hipoarticulação que

voluntários. Quanto à taxa de elocução, os resultados são divergentes: alguns estudos apontam redução significativa (Ludlow; Connor; Bassich, 1987), enquanto outros mostram variações opostas (Ackermann; Konczak; Hertrich, 1997). Patel, Parveen e Anand (2016) analisaram a prosódia em pessoas com DP, incluindo estrutura frasal, acento, taxa de elocução, frequência fundamental, intensidade e pausas. Os déficits de prosódia incluíram monotonia, volume constante e alterações na taxa de elocução e pausas.

A incoordenação pneumofonoarticulatória pode causar pausas imprevistas durante a fala, resultando em um padrão de fala irregular. Adicionalmente, a presença de fendas glóticas e tremor na DP contribuem para um aumento do esforço vocal. Apesar de haver alguma sobreposição nas características fônicas, a distinção, muitas vezes, reside nas causas subjacentes e nas manifestações clínicas. Estudos identificaram uma disritmia na fala de pessoas com DP, afetando a fluência e a relação fala/pausa (Lowit et al., 2018).

No grupo de idosos com DP, analisamos se há diferença significativa entre os resultados dos parâmetros prosódico-acústicos entre dois graus de severidade da DP e dois estilos de fala. O grau de severidade é medido pela Escala de Estágios de Incapacidade de Hoehn e Yahr (Hoehn; Yahr, 1967) de acompanhamento clínico longitudinal, aplicada por meio de um questionário que compõe a avaliação clínica e o diagnóstico diferencial realizado em atendimentos ambulatoriais de hospitais e clínicas especializadas em neurologia. Gravamos dois estilos de fala para cada indivíduo: a leitura de um parágrafo, seguido de um breve comentário espontâneo sobre o conteúdo lido.

Devido à DP, os indivíduos podem apresentar comprometimentos fônicos inclusive na fase prodromal<sup>3</sup> da doença. No segundo e terceiro estágios, as queixas vocais são relatadas com maior frequência (Braak et al., 2002; Greffard et al., 2006). A partir do segundo estágio da DP, ocorrem alterações no pitch, na intensidade, na frequência fundamental (F0), nas taxas de elocução e articulação, o que compromete a inteligibilidade da fala em ambos os sexos (Ackermann; Konczak; Hertrich, 1997). Ademais, cerca de 90% dos indivíduos com DP desenvolvem algum distúrbio vocal (Barbosa et al., 2006).

Os resultados deste estudo contribuem para melhor descrever os efeitos do envelhecimento nos parâmetros fônicos e fornecer subsídios para o aprimoramento de intervenções terapêuticas direcionadas à melhoria da comunicação e da qualidade de vida da população idosa com DP.

---

lentifica a produção sonora. A taxa de elocução considera o tempo total de fala, incluindo as pausas silenciosas, refletindo a cadência geral.

<sup>3</sup>Prodromal é um termo utilizado para se referir ao estágio inicial ou pré-sintomático de uma doença. É a fase que antecede o surgimento dos sintomas característicos de uma condição médica. Na DP, a fase prodromal se refere ao período antes do início dos sintomas motores clássicos, como tremores e rigidez muscular. Durante essa fase, algumas mudanças sutis podem ocorrer, como alterações na função olfativa, constipação, distúrbios do sono e outros sinais que podem indicar o desenvolvimento posterior da doença.

## 2 OBJETIVO

Este estudo tem como propósito analisar e sistematizar as diferenças nos parâmetros prosódico-acústicos entre um grupo experimental de idosos com DP, subdivididos em dois grupos de acordo com a severidade da doença, e um grupo controle de idosos saudáveis em dois estilos de fala: leitura e comentário espontâneo.

## 3 MÉTODO

Esta pesquisa delinea-se como um estudo descritivo, conduzido de acordo com os princípios éticos estabelecidos na Resolução nº 466/12 do Conselho Nacional de Saúde e foi aprovada pelo Comitê de Ética em Pesquisa da Universidade Estadual de Campinas (Parecer: 5.030.137). Os participantes receberam um Termo de Consentimento Livre e Esclarecido (TCLE) para assinatura, no qual foram informados sobre os objetivos, procedimentos, riscos e benefícios da pesquisa. O sigilo e a privacidade dos participantes foram preservados. Os participantes tinham entre 60 e 93 anos, eram falantes do português brasileiro e residentes do estado de São Paulo. Ademais, os idosos do grupo experimental tinham os laudos médicos da DP e as descrições dos sintomas nos seus prontuários. A ampla faixa etária contribui para a robustez dos resultados, oferecendo uma visão abrangente dos efeitos desses parâmetros nos dados de voz.

Foram analisadas gravações de dois grupos de idosos. O primeiro é um grupo controle com 16 idosos saudáveis. O segundo é o grupo experimental com 16 idosos diagnosticados com DP e subdivididos igualmente de acordo com o segundo e terceiro estágios da escala de Hoehn e Yahr. Os dois grupos e os dois subgrupos tinham o mesmo número de participantes de ambos os sexos.

Com o objetivo de selecionar os idosos no segundo e terceiro estágios da DP, foram considerados os sintomas descritos nos prontuários disponíveis no Hospital de Clínicas da Universidade Estadual de Campinas.

### 3.1 Participantes

Todos os participantes estavam em estado *on*<sup>4</sup> da medicação dopaminérgica (Armstrong; Okun, 2020) no momento da coleta, e, com o objetivo de caracterizar os participantes da pesquisa, foram compiladas informações em quadros, abrangendo idade, escolaridade, profissão e tempo de acompanhamento fonoaudiológico. Esses elementos se mostram pertinentes para a análise dos parâmetros prosódico-acústicos, pois a profissão pode sugerir o uso recorrente da voz, como é o caso de profissionais vocais (tais como cantores, professores, locutores e atores). A escolaridade dos participantes auxilia na consideração de possíveis diferenças na habilidade de leitura e comunicação vocal, mas esse fator não foi aqui considerado. O tempo de acompanhamento fonoaudiológico foi verificado no prontuário. Foram feitas

---

<sup>4</sup> Estado *on* na doença de Parkinson refere-se ao período em que os sintomas motores da doença estão sob controle devido à medicação. Durante esse estado, os pacientes geralmente apresentam, mesmo que temporariamente, uma melhora na mobilidade, diminuição do tremor, aumento da fluidez nos movimentos e melhora geral na funcionalidade. Isso ocorre em contraste com o estado *off*, quando os sintomas retornam ou se agravam devido ao declínio dos efeitos da medicação.

perguntas sobre os hábitos nocivos à voz, os mais significativos estão descritos nos quadros 1, 2 e 3.

A nomenclatura adotada para nomear os arquivos de áudio das gravações seguiu a sequência de critérios e ordem: 1. Indicação de Controle (CT) ou Experimental (DP); 2. Identificação de gênero Masculino (M) ou Feminino (F); 3. Letra para distinguir cada participante individualmente. A seguir, encontram-se os quadros com as informações relativas a idade, escolaridade, profissão e hábitos prejudiciais à voz dos participantes.

Quadro 1 – Grupo Controle de idosos em processo de envelhecimento saudável.

Participante	Idade	Escolaridade	Profissão	Hábitos nocivos à voz ao longo da vida
CTFC	61	Superior Completo	Advogada	Nenhum relatado
CTFG	79	Superior Completo	Assistente Social	Nenhum relatado
CTFL	68	Ensino Médio Completo	Costureira	Nenhum relatado
CTFM	62	Ensino Superior Completo	Atendente	Nenhum relatado
CTFO	66	Ensino Médio Completo	Recepcionista	Nenhum relatado
CTFR	62	Ensino Superior Completo	Professora de Ensino Fundamental	Nenhum relatado
CTFS	65	Ensino Superior Completo	Professora de Ensino Médio	Nenhum relatado
CTFI	67	Ensino Superior Completo	Professora de Ensino Superior	Nenhum relatado
CTMD	66	Ensino Médio Completo	Florista	Nenhum relatado
CTMC	60	Ensino Superior Completo	Técnico em Informática	Nenhum relatado
CTMG	62	Ensino Superior Completo	Auxiliar Administrativo	Nenhum relatado
CTMN	76	Ensino Superior Completo	Garimpeiro	Nenhum relatado
CTML	62	Ensino Médio Completo	Agricultor	Nenhum relatado
CTMD	62	Ensino Superior Completo	Auxiliar administrativo	Nenhum relatado
CTMT	62	Ensino Superior Completo	Engenheiro Civil	Nenhum relatado

CTMP	63	Ensino Superior Completo	Administrador de Empresas	Nenhum relatado
------	----	--------------------------	---------------------------	-----------------

Fonte: Elaborado pelos autores.

Quadro 2 – Grupo Experimental de idosos com DP no segundo estágio da escala de Hoehn e Yahr.

Participante	Idade	Escolaridade	Profissão	Hábitos nocivos à voz ao longo da vida	Tempo de acompanhamento fonoaudiológico
DPFC	65	Ensino Médio Completo	Dona de Casa	Nenhum relatado	6 anos e 2 meses
DPFI	66	Ensino Técnico Completo	Atendente	Nenhum relatado	3 anos e 0 meses
DPFJ	64	Ensino Fundamental Completo	Dona de Casa	Nenhum relatado	2 anos e 6 meses
DPFS	61	Ensino Superior Completo	Dona de Casa	Nenhum relatado	4 anos e 2 meses
DPMJ	70	Ensino Médio Completo	Garimpeiro	Nenhum relatado	2 anos e 3 meses
DPMK	73	Ensino Fundamental Completo	Vendedor	Fumante por 10 anos e parou há 5 anos	4 anos
DPMM	69	Ensino Básico Completo	Agricultor	Fumante por 5 anos e parou há 15 anos	2 anos e 6 meses
DPMR	70	Ensino Básico Completo	Vendedor	Nenhum relatado	3 anos

Fonte: Elaborado pelos autores.

Quadro 3 – Grupo Experimental de idosos com DP no terceiro estágio da escala de Hoehn e Yahr.

Participante	Idade	Escolaridade	Profissão	Hábitos nocivos à voz ao longo da vida	Tempo de acompanhamento fonoaudiológico
DPFY	71	Ensino Técnico Completo	Professora	Nenhum relatado	2 anos e 6 meses
DPFX	70	Ensino Médio Completo	Dona de Casa	Nenhum relatado	3 anos
DPFD	68	Ensino Médio Completo	Dona de Casa	Nenhum relatado	3 anos
DPFT	65	Ensino Superior Completo	Recepcionista	Nenhum relatado	4 anos
DPMA	93	Ensino Superior Completo	Advogado	Fumante por 20 anos	4 anos

DPMH	68	Ensino Superior Completo	Professor de Línguas	Nenhum relatado	2 anos e 6 meses
DPMN	70	Ensino Médio Completo	Garimpeiro	Nenhum relatado	1 ano e 6 meses
DPMW	66	Ensino Fundamental Completo	Vendedor	Nenhum relatado	2 anos

Fonte: Elaborado pelos autores.

Do grupo experimental, nove participantes estavam em acompanhamento pelo ambulatório de DP do Hospital de Clínicas da Universidade Estadual de Campinas e em atendimento fonoaudiológico no Centro de Estudos e Pesquisa em Reabilitação “Prof. Dr. Gabriel Porto”, enquanto os outros sete participantes desse grupo realizavam acompanhamentos particulares e já haviam sido avaliados no ambulatório de DP.

### 3.2 Forma de coleta de dados

O tempo de coleta de dados foi de 40 minutos para cada participante. Devido à pandemia da Covid-19, ocorrida no período de coleta dos dados, e os riscos associados a ela, os participantes foram informados sobre os riscos e que apenas realizaríamos as gravações presencialmente, caso se sentissem seguros. Nesse contexto, eles tiveram a opção de gravar usando um aplicativo gratuito em seus celulares com orientações remotas ou realizar a gravação na presença do pesquisador. Quatro participantes do grupo controle e quatro participantes do grupo experimental escolheram a opção de gravação remota, enquanto os demais optaram pela gravação presencial junto ao pesquisador.

Sobre a forma de gravação, os estudos de Maryn et al. (2017) indicam que diferentes dispositivos de gravação podem introduzir variações na medição da F0. Isso ocorre devido a diferenças nas características técnicas dos dispositivos, como resposta de frequência e sensibilidade. Além disso, o ambiente de gravação, incluindo o isolamento acústico, ruídos externos e reverberação, também podem afetar a precisão das medidas da F0. Por exemplo, gravações realizadas com um celular próprio podem apresentar uma resposta de frequência limitada ou um alto nível de ruído de fundo, o que pode resultar em imprecisões na medição da F0. Por outro lado, gravações feitas em uma cabine acústica projetada para minimizar a influência de ruídos e reflexões indesejadas geralmente produzem resultados mais precisos na análise da F0.

O sinal de áudio foi gravado onde estava o participante. O nível de ruído, de um modo geral, não comprometeu as medidas feitas. Para verificar a influência do ruído sobre a qualidade das medidas acústicas, inspecionamos o traçado de F0 gerado pelo Praat versão 6.4.07 (Boersma; Weenink, 2024) pela busca nos limites que dão a melhor continuidade do traçado, sem saltos bruscos. Não foram encontradas alterações que impactassem os descritores aqui usados. Segundo Maryn e colegas (2017), os parâmetros *shimmer*, *jitter* e relação harmônico-ruído (HNR) podem ser influenciados pelo ruído ambiente, comprometendo a precisão das análises acústicas. Além disso, consideramos que a ênfase espectral (*emph*) e o coeficiente de variação de intensidade (*cvint*) também podem ser afetados. No entanto, para avaliar o impacto do ruído de ambiente nesses parâmetros, analisamos a influência das gravações com

menores valores da relação sinal-ruído (SNR) nos resultados do cálculo do tamanho de efeito<sup>5</sup>.

Este estudo foi realizado a partir de amostras de leitura de texto e de comentários espontâneos. Esses estilos foram selecionados pela sua relevância na comunicação cotidiana e na avaliação clínica da fala. A leitura proporciona uma análise controlada da prosódia, enquanto o comentário espontâneo reflete a fala em um contexto próximo ao da conversa, capturando nuances que variam com o contexto e auxiliando na compreensão das alterações prosódicas no envelhecimento e na DP.

Os participantes foram solicitados a ler um parágrafo de 68 palavras. Em seguida, fizeram ao menos um comentário espontâneo sobre o trecho lido. A escolha do seguinte parágrafo se deve ao fato de abordar um tema que desperta familiaridade e memórias culturais relacionadas à alimentação de brasileiros nativos. Nesse contexto, a paçoca serve como uma referência cultural com o objetivo de evocar experiências pessoais dos participantes, como lembranças de sabores, tradições culinárias da família ou momentos compartilhados que tiveram a presença desse alimento. Segue o parágrafo adaptado de uma publicação do blog de uma empresa de produtos alimentícios (WEF, 2020) e embasado pelos estudos de Cascudo (2004, p. 157):

A paçoca servia como alimento dos garimpeiros que saíam para procurar alimentos às margens dos rios. A paçoca era feita de carne e com o passar dos anos ela foi aperfeiçoada com outros ingredientes. O nome paçoca vem de um termo indígena que quer dizer “esmagar com as mãos”, referindo-se a um modo antigo de preparo em que a carne e a farinha eram amassadas em um pilão.

O parágrafo apresenta informações sobre a origem e o processo de preparo da paçoca, o que pode despertar curiosidade e levar os participantes a compartilharem seus conhecimentos sobre o assunto espontaneamente. Após a gravação, os áudios foram analisados com uma metodologia específica que será apresentada na próxima seção.

---

<sup>5</sup> Quatro gravações do grupo controle — duas do grupo feminino e duas do grupo masculino — apresentaram valores de SNR de 22,1 dB e 23,92 dB no grupo masculino, e 24,87 dB e 26,38 dB no grupo feminino. Para avaliar o impacto dessas gravações nos parâmetros *jitter*, *shimmer*, HNR, *cvint* e *emph*, excluímos essas gravações e repetimos os testes SRH em ambos os sexos. No grupo masculino, a ordem dos cinco parâmetros com maior tamanho de efeito permaneceu inalterada. No grupo feminino, a ordem decrescente, antes de retirar os áudios com SNR < 30 dB era: *cvint* (primeiro), *emph* (segundo), HNR (nono). Depois de retirar as duas gravações já referidas do grupo feminino, a ordem mudou para: *emph* (primeiro), HNR (terceiro) e *cvint* (quinto). Os parâmetros derivados da F0 são resistentes aos efeitos do ruído ambiental (Maryn et al., 2017, p. 248.e11), razão pela qual não atribuímos quaisquer alterações nesses parâmetros aos efeitos do ruído nas duas gravações do grupo feminino. Adicionalmente, os parâmetros de taxa de elocução e taxa de articulação, por serem resultados do número de sílabas produzidas em trechos de fala, e *sLTASmed*, por utilizar a mediana e a suavização do LTAS, são menos influenciados pelo ruído ambiental. Os parâmetros *emph* e *cvint* se mantiveram entre os cinco com maior tamanho de efeito e, embora o HNR tenha apresentado maior tamanho após a retirada das já referidas gravações, esse parâmetro apresentou consistência no nível de significância, assim como os demais parâmetros sensíveis ao ruído ambiental: *emph*, *cvint*, *jitter* e *shimmer*. Isso confere robustez aos resultados obtidos no grupo feminino.

### 3.3 Forma de análise dos resultados

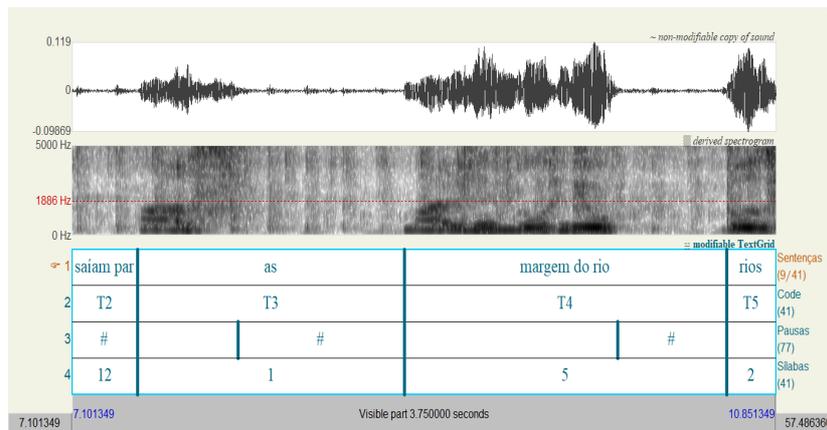
A análise acústica das gravações foi realizada utilizando um script desenvolvido por Barbosa (2023) para o programa Praat versão 6.4.07 (Boersma; Weenink, 2024), chamado Prosody Descriptor Extractor<sup>6</sup>. Os valores extraídos foram analisados estatisticamente no software R. O script permite a análise das seguintes classes de parâmetros acústicos, a partir das quais foram extraídos descritores estatísticos em trechos delimitados por pausas silenciosas ou preenchidas em uma camada, e pelo número de sílabas fonológicas em uma terceira camada (cf. Figura 1): i. F0 - inclui média, mediana, desvio-padrão, largura de banda e taxa de picos por segundo; ii. taxa de variação da F0 - considera a primeira derivada da F0, analisando separadamente as subidas e descidas e seus respectivos valores médios e desvios-padrão; iii. intensidade relativa - ênfase espectral, definida como a intensidade total menos a intensidade até 400 Hz (Traunmüller; Eriksson, 2000); iv. coeficiente de variação da intensidade - razão entre o desvio-padrão e a média da intensidade<sup>7</sup>, expressa em porcentagem; v. inclinação do LTAS (*long-term average spectrum*)<sup>8</sup> - inclinação do LTAS entre as bandas 0-1000 Hz e 1000-4000 Hz (sLTASmed) e inclinação do LTAS entre as bandas 0-1000 Hz e 4000-8000 Hz (sLTAShigh), conforme descrito por Traunmüller e Eriksson (2000), Dromey (2003) e Tanner et al. (2005); e vi. relação harmônico-ruído (*Harmonic-to-Noise Ratio -hnr*), *jitter* e *shimmer*. Além disso, o script permite a análise de outros parâmetros, como a frequência de ocorrência e a duração das pausas, bem como as taxas de elocução e de articulação.

Para a análise dos parâmetros, segmentamos as amostras de fala da seguinte forma: trecho a trecho das gravações nos dois estilos (leitura e comentário), com auxílio do espectrograma, marcação manual do número de sílabas fonológicas (unidades compostas por núcleo, ataque e coda) e de pausas silenciosas e preenchidas (em milissegundos). As fronteiras das pausas silenciosas foram determinadas pela primeira marcação corresponde ao final da última vogal do segmento emitido antes da pausa, e a marcação inicial corresponde ao início da primeira vogal do segmento seguinte. A razão pela escolha de tal marcação de pausa fundamenta-se nos trabalhos de Grosjean e Lane (1977) e Monnin e Grosjean (1993), pois melhor revela a hierarquia das forças prosódicas subjacentes, o que é confirmado no Português Brasileiro (PB) no trabalho de Barbosa (1999). Segue um exemplo da forma de segmentação.

<sup>6</sup> O Prosody Descriptor Extractor é um script disponível online e fornece até 25 descritores prosódicos para enunciados completos ou trechos do mesmo enunciado, com o propósito de viabilizar a pesquisa sobre a relação entre a produção e a percepção da prosódia. Este script possui como parâmetros de entrada os nomes dos arquivos de Som e TextGrid correspondentes. Disponível em: <https://github.com/pabarbosa/prosodyscripts/tree/master/ProsodyDescriptorExtractor>.

<sup>7</sup> Como a medida da intensidade é uma razão, ela é imune à distância do microfone.

<sup>8</sup> LTAS é um espectro de potência gerado por transformada rápida de Fourier das frequências que compõem uma amostra de fala. Portanto, o LTAS é um sinal composto que representa o espectro da fonte glotal, bem como o espectro ou as características ressonantes do trato vocal. O LTAS apresenta potencial como um índice acústico de qualidade vocal, pelo menos quando a produção dos segmentos está razoavelmente intacta. Por exemplo, uma energia harmônica relativamente fraca nas frequências mais altas do espectro da fala e um aumento correspondente na inclinação espectral são características de sinais sussurrados ou hipofuncionais. Em contraste, o impacto excessivo das pregas vocais e o ruído turbulento, ambos observados na disfonia funcional, estão associados a uma energia relativamente maior nas frequências mais altas do espectro da fala (Dromey, 2003; Tanner et al., 2005).



Fonte: Elaborada pelos autores.

Figura 1 – Trechos da fala do participante do grupo experimental delimitados por pausas preenchidas ou silenciosas (#) e número de sílabas fonológicas (última camada).

### 3.4 Análise estatística

A análise estatística foi conduzida usando o software RStudio versão 2023.06.1 build 524 (RStudio Team, 2023) e envolveu a avaliação da normalidade dos dados por meio do teste de *Shapiro-Wilk* ( $p < 0,05$ ) (Pagano; Gauvreau, 2018). Devido à ausência de normalidade e à heterocedasticidade dos resíduos originais, optamos por empregar a técnica estatística conhecida como extensão *Scheirer-Ray-Hare* (SRH) ao teste de *Kruskal-Wallis* (Dytham, 2011). Essa abordagem oferece uma alternativa não paramétrica à ANOVA de dois fatores, levando em consideração a interação entre os fatores e suas diferenças de classificação em cada nível, com base na soma dos quadrados das discrepâncias entre as classificações observadas e esperadas para cada nível de cada fator.

Para identificar as diferenças significativas reveladas pelo teste não paramétrico SRH, recorreremos ao teste não paramétrico de *Wilcoxon* para comparações em pares. Esse teste é utilizado para realizar comparações múltiplas entre todas as combinações possíveis de pares de grupos em um conjunto de dados. O teste de *Wilcoxon* gera uma matriz triangular de valores-p, os quais são ajustados pelo método de correção de Bonferroni.

Os graus de liberdade foram calculados com base no número de fatores e de seus níveis. Consideramos que os fatores *grupo*, *severidade* e *estilo* desempenharam papéis essenciais nos resultados da análise de variância. O fator de *severidade* da DP organiza a amostra em três níveis: um nível para o grupo controle (SC) e os estágios da escala Hoehn e Yahr da DP, um nível para o estágio 2 (DP2) e um nível para o estágio 3 (DP3); já o fator *estilo de gravação* organiza a amostra em dois níveis: um nível para o estilo comentário (C) e outro para o estilo leitura (L).

Analizamos as diferenças médias no contexto de múltiplos níveis do fator *estilo* e investigamos as combinações de níveis com o fator *severidade* para identificar diferenças significativas entre os grupos deste estudo. Não foram avaliadas diferenças entre os participantes quanto aos níveis de escolaridade. Assim, para a análise das produções, usamos o teste SRH em quatro modelos:

- (1) *estilo* e *grupo* para o grupo feminino;

- (2) *estilo* e *severidade* para o grupo feminino;
- (3) *estilo* e *grupo* para o grupo masculino;
- (4) *estilo* e *severidade* para o grupo masculino.

Devido ao grande número de parâmetros acústicos que analisamos, 24, calculamos o tamanho do efeito das variáveis dependentes (parâmetros prosódico-acústicos) em relação ao fator *estilo* em cada sexo e geramos gráficos utilizando a biblioteca *ggplot2* no RStudio versão 2023.06.1 (RStudio Team, 2023) com os resultados significativos. Na análise de variância SRH, identificamos quatro fontes de variação que explicam os resultados da variável dependente: *primeiro fator*, *segundo fator*, *interação entre fatores* e *resíduos*. O tamanho do efeito foi calculado por meio do coeficiente de eta quadrado ( $\eta^2$ ), que representa a proporção da variância total explicada pelo fator:

$$(5) \eta^2 = \frac{\text{Soma das Quadrados dos erros devido ao fator}}{\text{Soma Total dos Quadrados dos erros}}$$

Essa fórmula resulta em um número que pode variar de 0 a 1 e indica a proporção da variância total do parâmetro-acústico que é explicado pelo fator *estilo*.

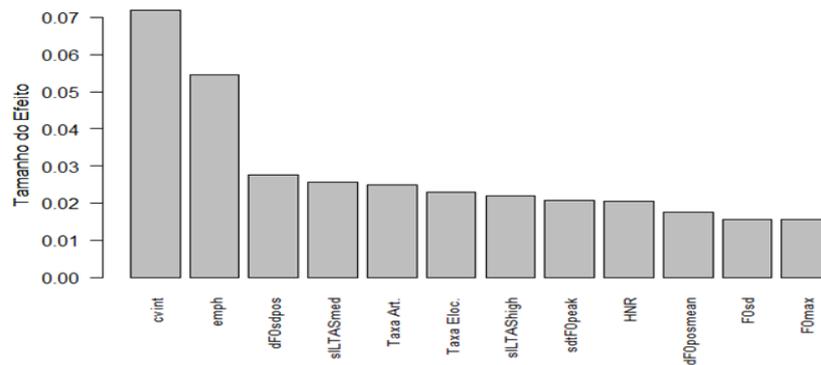
Na próxima seção, apresentaremos os resultados dos modelos de SRH e os valores *p* dos parâmetros prosódico-acústicos, de acordo com o nível de significância <0,05, e o tamanho do efeito dos parâmetros em ordem de relevância por sexo.

#### 4 RESULTADOS

Apresentaremos primeiramente os resultados do cálculo do tamanho do efeito nos modelos SRH 2-Fatores, seguido dos parâmetros com maior tamanho do efeito nos modelos. Quanto à análise estatística dos dados acústicos, foi aplicado o teste não paramétrico SRH com todos os fatores, mostrando efeitos principais e interações significativas. As seções a seguir apresentam os parâmetros com os maiores tamanhos de efeito, juntamente com os resultados dos testes *post hoc* de *Wilcoxon* aplicando a correção de Bonferroni. Em todos os casos,  $\alpha = 0,05$ .

Para os modelos SRH no grupo feminino, tanto o que considera o estilo de fala em relação aos grupos controle e experimental quanto o modelo que leva em consideração o estilo de fala em relação aos grupos controle, segundo estágio da DP e terceiro estágio da DP, os valores *p* dos seguintes parâmetros prosódico-acústicos foram significativos para o fator *estilo* em ordem decrescente de tamanho do efeito: coeficiente de variação de intensidade (*cvint*), ênfase espectral (*emph*), desvio padrão da primeira derivada positiva da F0 (*df0sdpos*), inclinação de LTAS em frequências médias (*sLTASmed*), taxa de articulação, taxa de elocução, inclinação de LTAS em frequência altas (*sLTAShigh*), desvio padrão das posições dos picos da F0 (*sdtF0peak*), relação harmônico-ruído (HNR), média da primeira derivada positiva da F0 (*df0posmean*), desvio padrão da F0 (*F0sd*) e máximo da F0 (*F0max*). Segue o Gráfico 1<sup>9</sup> que representa o tamanho do efeito dos parâmetros no grupo feminino.

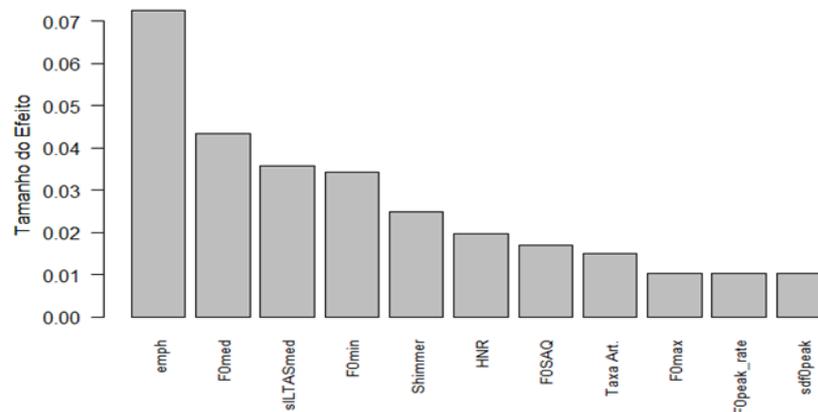
<sup>9</sup> O eixo y representa o tamanho do efeito dos parâmetros prosódico-acústicos, considerando o fator *estilo* em relação aos grupos *controle* e *estágios da DP*, enquanto o eixo x representa os diferentes parâmetros. Cada barra representa um parâmetro, ordenado pelo tamanho do efeito.



Fonte: Elaborado pelos autores.

Gráfico 1 – Tamanho do efeito no primeiro e segundo modelos SRH – feminino.

Para os modelos SRH no grupo masculino, tanto o modelo que considera o estilo de fala em relação aos grupos controle e experimental quanto o modelo que leva em consideração o estilo de fala em relação aos grupos controle, segundo estágio da DP e terceiro estágio da DP, os valores *p* dos seguintes parâmetros prosódico-acústicos foram significativos para o fator *estilo* em ordem decrescente de tamanho do efeito: ênfase espectral (*emph*), mediana da F0 (*F0med*), inclinação de LTAS em frequência média (*sILTASmed*), mínimo da F0 (*F0min*), *shimmer*, relação harmônico-ruído (*HNR*), semi-amplitude entre os quartis da F0 (*F0SAQ*), desvio padrão da primeira derivada negativa da F0 (*dF0sdneg*), taxa de articulação, máximo da F0 (*F0max*), taxa de picos da F0 (*F0peakrate*) e desvio padrão das posições dos picos da F0 (*sdF0peak*). Segue o Gráfico 2<sup>10</sup>, no qual aparece representado o tamanho do efeito dos parâmetros no grupo masculino.



Fonte: Elaborado pelos autores.

Gráfico 2 – Tamanho do efeito no terceiro e quarto modelos SRH – masculino.

O idoso mais velho da amostra, com 93 anos e no terceiro estágio da DP, apresenta diferenças significativas nos estilos de fala em relação ao grupo de homens no mesmo estágio nos parâmetros *F0med*, *F0min*, *cvint* e *jitter*. Para os dois estilos de

<sup>10</sup> O eixo y representa o tamanho do efeito dos parâmetros prosódico-acústicos, considerando o fator *estilo* em relação aos grupos controle e estágios da DP, enquanto o eixo x representa os diferentes parâmetros. Cada barra representa um parâmetro, ordenado pelo tamanho do efeito.

fala em relação ao grupo DP3, o idoso mais velho apresentou médias mais baixas para F0med e F0min e médias maiores para cvint e *jitter*.

A ordem dos parâmetros com maior tamanho de efeito da amostra, excluindo o participante mais velho, varia, mas os cinco primeiros são consistentemente os mesmos para ambos os modelos SHR 2-Fatores: *emph* (0,09), F0med (0,05), *shimmer* (0,041), sLTASmed (0,04), F0min (0,038).

A seguir, apresentamos os parâmetros com maior tamanho de efeito na amostra que usamos para analisar as variáveis dependentes (parâmetros prosódico-acústicos).

#### 4.1 Parâmetros com maior tamanho de efeito

Selecionamos os cinco parâmetros com o maior tamanho de efeito em cada modelo SRH. O primeiro e o segundo modelos compartilharam os mesmos cinco parâmetros de maior tamanho de efeito e em ordem de relevância, assim como os terceiro e quarto modelos.

Nos primeiro e segundo modelos, os parâmetros com maior efeito em ordem de relevância foram: coeficiente de variação da intensidade (cvint), ênfase espectral (*emph*), desvio padrão da primeira derivada positiva da F0 (dF0sdpos), inclinação de LTAS em frequência média (sLTASmed) e taxa de articulação.

Nos terceiro e quarto modelos, os parâmetros com maior efeito em ordem de relevância foram: ênfase espectral (*emph*), mediana da F0 (F0med), inclinação de LTAS em frequência média (sLTASmed), mínimo da F0 (F0min) e *shimmer*. A seguir, apresentamos os resultados desses parâmetros prosódico-acústicos com maior efeito na variância total. Na próxima seção, apresentamos os resultados de acordo com o sexo.

#### 4.2 Parâmetros significativos para ambos os sexos

A ênfase espectral foi o segundo parâmetro de maior impacto na variância total entre as mulheres. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo feminino para os fatores de estilo de fala, comentário (C) e leitura (L), e estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 1 - *Post-hoc* não paramétrico da *emph* para *estilo* e *severidade* do grupo feminino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1,00	-	-	-	-
DP3.C	1,00	1,00	-	-	-
DP3.L	1x10 <sup>-3</sup>	9x10 <sup>-4</sup>	1x10 <sup>-3</sup>	-	-
CT.C	0,03	6x10 <sup>-6</sup>	0,01	4x10 <sup>-14</sup>	-
CT.L	2x10 <sup>-3</sup>	1x10 <sup>-7</sup>	2x10 <sup>-3</sup>	<2x10 <sup>-16</sup>	1,00

Fonte: Elaborada pelos autores.

Na Tabela 1, identificamos diferenças significativas na ênfase espectral entre os estágios de severidade: controle-comentário e DP2-comentário ( $p = 0,03$ ), DP2-leitura e DP3-leitura ( $p = 9 \times 10^{-4}$ ), controle-leitura e DP2-leitura ( $p = 1 \times 10^{-7}$ ), controle-comentário e DP3-comentário ( $p = 0,01$ ) e controle-leitura e DP3-leitura ( $p = < 2 \times 10^{-16}$ ). Ao comparar os estilos de fala, somente os grupos DP3-leitura e DP3-comentário ( $p = 1 \times 10^{-3}$ ) se diferenciaram.

Os valores médios da ênfase espectral do grupo feminino para os fatores de estilo de fala de acordo com os estágios de severidade da DP foram: 1,2 dB em DP3-leitura, 1,8 dB em DP2-leitura, 2,3 dB em DP2-comentário, 2,5 dB em DP3-comentário, 3,7 dB em controle-leitura e 4 dB em controle-comentário

No grupo masculino, a ênfase espectral foi o parâmetro com maior efeito na variância total. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo masculino para os fatores de estilo de fala, comentário (C) e leitura (L), e estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 2 - *Post-hoc* não paramétrico da *emph* para *estilo e severidade* do grupo Masculino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	0.03	-	-	-	-
DP3.C	0.95	5x10 <sup>-10</sup>	-	-	-
DP3.L	1.00	6x10 <sup>-4</sup>	0.19	-	-
CT.C	1.00	7x10 <sup>-5</sup>	0.10	1.00	-
CT.L	1.00	0.03	2x10 <sup>-5</sup>	0.68	0.76

Fonte: Elaborada pelos autores.

Na Tabela 2, identificamos diferenças significativas na ênfase espectral entre os estágios de severidade: DP3-leitura e DP2-leitura ( $p = 6 \times 10^{-4}$ ) e controle-leitura e DP2-leitura ( $p = 0,03$ ). Ao compararmos os estilos de fala, somente os grupos DP2-leitura e DP2-comentário ( $p = 0,03$ ) se diferenciaram.

Os valores médios da ênfase espectral do grupo masculino para os fatores de estilo de fala, de acordo com os estágios de severidade da DP, foram: 2,5 dB no DP3-comentário, 2 dB no DP3-leitura, 2 dB no controle-comentário e 2 dB no DP2-comentário, 1,5 dB no controle-leitura e 1,1 dB no DP2-leitura.

O parâmetro *sLTASmed* teve o quarto maior tamanho de efeito no grupo feminino. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo feminino para os fatores de estilo de fala, comentário (C) e leitura (L), estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 3 - *Post-hoc* não paramétrico da *sLTASmed* para *estilo e severidade* do grupo feminino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1.00	-	-	-	-
DP3.C	0.75	0.38	-	-	-
DP3.L	1x10 <sup>-4</sup>	2x10 <sup>-8</sup>	3x10 <sup>-3</sup>	-	-
CT.C	1.00	0.22	2x10 <sup>-3</sup>	7x10 <sup>-12</sup>	-
CT.L	0.67	0.03	1x10 <sup>-4</sup>	5x10 <sup>-16</sup>	1.00

Fonte: Elaborada pelos autores.

Na Tabela 3, identificamos diferenças significativas na inclinação de LTAS em frequência média entre os estágios de severidade: DP3-leitura e DP2-leitura ( $p = 2 \times 10^{-8}$ ), controle-leitura e DP2-leitura ( $p = 0,03$ ), controle-comentário e DP3-comentário ( $p = 2 \times 10^{-3}$ ) e controle-leitura e DP3-leitura ( $p = 5 \times 10^{-16}$ ). Ao comparar os estilos de fala, somente os grupos DP3-leitura e DP3-comentário ( $p = 3 \times 10^{-3}$ ) se diferenciaram.

Os valores médios da inclinação de LTAS em frequência média do grupo feminino para os fatores de estilo de fala e estágios de severidade da DP foram: -29,5 dB/quadro no DP3-leitura, -21 dB/quadro DP3-comentário, -18 dB/quadro no DP2-comentário, -18 dB/quadro no DP2-leitura, -16,5 dB/quadro no controle-comentário, -15,5 dB/quadro no controle-leitura.

O sLTASmed foi o terceiro parâmetro com maior impacto na variância total do grupo masculino. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo masculino para os fatores de estilo de fala, comentário (C) e leitura (L), estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 4 - *Post-hoc* não paramétrico do sLTASmed para *estilo e severidade* do grupo masculino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1.00	-	-	-	-
DP3.C	$4 \times 10^{-6}$	$2 \times 10^{-13}$	-	-	-
DP3.L	0.07	$4 \times 10^{-6}$	0.07	-	-
CT.C	1.00	0.03	$1 \times 10^{-6}$	0.17	-
CT.L	1.00	0.06	$4 \times 10^{-7}$	0.10	1.00

Fonte: Elaborada pelos autores.

Na Tabela 4, identificamos diferenças significativas na inclinação de LTAS em frequência média entre os estágios de severidade: DP2-comentário e DP3-comentário ( $p = 4 \times 10^{-6}$ ), DP3-leitura e DP2-leitura ( $p = 4 \times 10^{-6}$ ), controle-leitura e DP2-leitura ( $p = 0,03$ ) e DP3-comentário e controle-comentário ( $p = 1 \times 10^{-6}$ ). No entanto, não foram encontradas diferenças significativas entre os estilos de fala.

Os valores médios da inclinação de LTAS em frequência média do grupo masculino para os fatores de estilo de fala e estágios de severidade da DP foram: -17,5 dB/quadro no DP3-comentário, -20 dB/quadro no DP3-leitura, -21 dB/quadro no controle-comentário, -22 dB/quadro no controle-leitura, -22,5 dB/quadro no DP2-comentário e -24,5 dB/quadro DP2-leitura.

#### 4.2 Parâmetros significativos para o sexo feminino

O coeficiente de variação de intensidade foi o parâmetro de maior tamanho de efeito para as mulheres. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo feminino para os fatores de estilo de fala, comentário (C) e leitura (L), estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 5 - *Post-hoc* não paramétrico do cvint para *estilo e severidade* do grupo feminino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1.00	-	-	-	-
DP3.C	1.00	1.00	-	-	-
DP3.L	0.15	0.90	0.56	-	-
CT.C	0.25	$5 \times 10^{-7}$	0.26	$2 \times 10^{-9}$	-
CT.L	1.00	0.04	1.00	$5 \times 10^{-5}$	0.06

Fonte: Elaborada pelos autores.

Na Tabela 5, identificamos diferença significativa no coeficiente de variação de intensidade entre os estágios controle-leitura e DP3-leitura ( $p = 5 \times 10^{-5}$ ). No entanto, não foram encontradas diferenças significativas entre os estilos de fala.

Os valores médios do coeficiente de variação da intensidade do grupo feminino para os fatores de estilo de fala e estágios de severidade da DP foram: 20% no DP3-leitura, 18,26% no DP2-leitura, 17,37% no DP3-comentário, 16,2% DP2-comentário, 15,76% no controle-leitura e 13,64% no controle-comentário.

O desvio padrão da primeira derivada positiva da frequência fundamental foi o terceiro parâmetro com maior efeito no grupo feminino. A seguir, apresentamos o

teste *post-hoc* não paramétrico do grupo feminino para os fatores de estilo de fala, comentário (C) e leitura (L), estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 6 - *Post-hoc* não paramétrico do  $dF_0$ sdpos para *estilo e severidade* do grupo feminino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1.00	-	-	-	-
DP3.C	1.00	1.00	-	-	-
DP3.L	0.13	0.80	0.01	-	-
CT.C	1.00	0.24	1.00	$1 \times 10^{-4}$	-
CT.L	1.00	0.04	1.00	$2 \times 10^{-6}$	1.00

Fonte: Elaborada pelos autores.

Na Tabela 6, identificamos diferenças significativas no desvio padrão da primeira derivada positiva da frequência fundamental entre os estágios de severidade: DP2-leitura e controle-leitura ( $p = 0,04$ ) e DP3-leitura e controle-leitura ( $p = 2 \times 10^{-6}$ ). No estilo de fala, somente os grupos DP3-leitura e DP3-comentário ( $p = 0,01$ ) se diferenciaram.

Os valores médios do desvio padrão da primeira derivada positiva da frequência fundamental do grupo feminino para os fatores de estilo de fala e estágios de severidade da DP foram: 2 Hz/quadro no DP3-leitura, 2,7 Hz/quadro no DP2-leitura, 3,2 Hz/quadro no DP3-comentário, 3,4 Hz/quadro no controle-leitura, 3,4 dB/quadro no DP2-comentário e 3,6 Hz/quadro no controle-comentário.

A taxa de articulação foi o quinto parâmetro com maior tamanho de efeito entre as mulheres. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo feminino para os fatores de estilo de fala, comentário (C) e leitura (L), estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 7 - *Post-hoc* não paramétrico da taxa de articulação para *estilo e severidade* do grupo feminino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1.00	-	-	-	-
DP3.C	1.00	1.00	-	-	-
DP3.L	0.84	1.00	0.62	-	-
CT.C	0.74	$3 \times 10^{-3}$	0.75	$1 \times 10^{-7}$	-
CT.L	0.20	$3 \times 10^{-5}$	0.14	$2 \times 10^{-11}$	1.00

Fonte: Elaborada pelos autores.

Na Tabela 7, identificamos diferenças significativas na taxa de articulação entre os estágios de severidade controle-leitura e DP2-leitura ( $p = 3 \times 10^{-5}$ ) e controle-leitura e DP3-leitura ( $p = 2 \times 10^{-11}$ ). No entanto, não foram encontradas diferenças significativas entre os estilos de fala.

Os valores médios da taxa de articulação do grupo feminino para os fatores de estilo de fala e estágios de severidade da DP foram: 3,5 sílabas/segundo no DP3-leitura, 4 sílabas/segundo no DP2-leitura, 4 sílabas/segundo DP2-comentário e 4,2 sílabas/segundo no DP3-comentário, 5 sílabas/segundo no controle-comentário e 5 sílabas/segundo no controle-leitura.

### 4.3 Parâmetros significativos para o sexo masculino

O parâmetro F0med foi o segundo com maior efeito no grupo masculino. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo masculino para os fatores de estilo de fala, comentário (C) e leitura (L), estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 8 - *Post-hoc* não paramétrico do F0med para *estilo e severidade* do grupo masculino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	0.44	-	-	-	-
DP3.C	1.00	0.02	-	-	-
DP3.L	1.00	1.00	0.05	-	-
CT.C	0.38	1.00	0.03	1.00	-
CT.L	$1 \times 10^{-5}$	$2 \times 10^{-5}$	$3 \times 10^{-10}$	$2 \times 10^{-5}$	0.03

Fonte: Elaborada pelos autores.

Na Tabela 8, identificamos diferenças significativas na mediana da frequência fundamental para os estágios de severidade da DP: DP3-leitura e controle-leitura ( $p = 2 \times 10^{-5}$ ), controle-comentário e DP3-comentário ( $p = 0,03$ ) e controle-leitura e DP2-leitura ( $p = 2 \times 10^{-5}$ ). Para o estilo de fala, há diferença somente entre controle-comentário e controle-leitura ( $p = 0,03$ ).

Os valores médios da mediana da frequência fundamental do grupo masculino para os fatores de estilo de fala e estágios de severidade da DP foram: 150,5 Hz no DP2-comentário, 149 Hz no DP3-comentário, 141 Hz no DP3-leitura, 141 Hz no controle-comentário, 140 Hz no DP2-leitura e 126 Hz no controle-leitura.

F0min foi o quarto parâmetro com maior efeito no grupo masculino. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo masculino para os fatores de estilo de fala, comentário (C) e leitura (L), e estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 9 - *Post-hoc* não paramétrico do F0min para *estilo e severidade* do grupo masculino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1.00	-	-	-	-
DP3.C	0.58	$2 \times 10^{-3}$	-	-	-
DP3.L	1.00	1.00	0.09	-	-
CT.C	1.00	1.00	$1 \times 10^{-3}$	1.00	-
CT.L	0.01	$1 \times 10^{-5}$	$6 \times 10^{-12}$	$4 \times 10^{-6}$	$2 \times 10^{-3}$

Fonte: Elaborada pelos autores.

Na Tabela 9, identificamos diferenças significativas na mediana da frequência fundamental entre os estágios de severidade: DP2-leitura e controle-leitura ( $p = 1 \times 10^{-5}$ ), controle-comentário e DP3-comentário ( $p = 0,03$ ), controle-comentário e DP3-comentário ( $p = 1 \times 10^{-3}$ ) e DP3-leitura e controle-leitura ( $p = 4 \times 10^{-6}$ ). Enquanto para os estilos de fala, somente os grupos controle-leitura e controle-comentário ( $p = 2 \times 10^{-3}$ ) se diferenciaram.

Os valores médios do mínimo da frequência fundamental do grupo masculino para os fatores de estilo de fala e estágios de severidade da DP foram: 132 Hz no DP3-comentário, 126 Hz no controle-leitura, 125 Hz no DP3-leitura, 124,5 Hz no DP2-comentário, 122 Hz no controle-comentário e 122 Hz no DP2-leitura.

O parâmetro *shimmer* foi o quinto com maior tamanho de efeito na amostra entre os homens. A seguir, apresentamos o teste *post-hoc* não paramétrico do grupo masculino para os fatores de estilo de fala, comentário (C) e leitura (L), e estágio de severidade da DP, controle (CT), segundo estágio (DP2) e estágio 3 (DP3).

Tabela 10 - *Post-hoc* não paramétrico do *shimmer* para *estilo e severidade* do grupo masculino.

	DP2.C	DP2.L	DP3.C	DP3.L	CT.C
DP2.L	1.00	-	-	-	-
DP3.C	0.36	0.07	-	-	-
DP3.L	1.00	1.00	0.02	-	-
CT.C	1.00	1.00	$9 \times 10^{-3}$	$3 \times 10^{-7}$	-
CT.L	0.01	$6 \times 10^{-4}$	$3 \times 10^{-13}$	$1 \times 10^{-6}$	$5 \times 10^{-4}$

Fonte: Elaboração própria.

Na Tabela 10, identificamos diferenças significativas em *shimmer* entre os grupos de severidade da DP: controle-leitura e DP2-leitura ( $p = 6 \times 10^{-4}$ ), controle-comentário e DP3-comentário ( $p = 9 \times 10^{-3}$ ) e controle-leitura e DP3-leitura ( $p = 1 \times 10^{-6}$ ). Para o estilo de fala, se diferenciaram: DP3-leitura e DP3-comentário ( $p = 0,02$ ) e controle-leitura e controle-comentário ( $5 \times 10^{-4}$ ).

Os valores médios de *shimmer* do grupo masculino para os fatores de estilo de fala e estágios de severidade da DP foram: 17% no DP3-comentário, 16,3% no DP3-leitura, 15,5% no DP2-leitura, 15,2% no DP2-comentário, 15% no controle-comentário e 13% no controle-leitura.

## 5 DISCUSSÃO

Dos 24 parâmetros prosódico-acústicos dos 32 idosos desse estudo, oito demonstraram capacidade de discriminação entre os grupos e subgrupos de idosos saudáveis e idosos com DP, destacando-se como aqueles que contribuíram significativamente para explicar a variância total da amostra. A seguir discutimos os resultados desses parâmetros.

Devido às restrições motoras da DP, a ênfase espectral (*emph*) pode ser afetada, pois está diretamente ligada ao sistema respiratório. Isso reflete o esforço vocal e a habilidade do falante em regular a saída de ar durante a fonação (Vieira, 2017). No processo de envelhecimento não patológico, as mudanças fisiológicas também podem influenciar na ênfase espectral, resultando na diminuição da elasticidade das estruturas vocais, da redução da massa muscular ao redor da laringe e vias respiratórias, e da atrofia das pregas vocais, o que contribui para alterações na qualidade vocal (Quedas et al., 2007).

O grupo feminino apresentou diferenças entre os estágios de severidade da DP e para os estilos de fala no terceiro estágio DP. Entre as mulheres, o subgrupo DP3-leitura teve o menor valor médio. O grupo masculino também revelou diferenças entre os estágios de severidade da DP; para os estilos de fala, o subgrupo DP2 se diferenciou e teve o menor valor médio do parâmetro durante a leitura.

Na DP, os parâmetros de qualidade vocal podem ser afetados devido à rigidez muscular, que influencia a vibração das pregas vocais e pode reduzir a área de fechamento glotal, resultando em uma maior perda de energia durante a fala. Isso, por sua vez, tem o potencial de alterar o parâmetro de inclinação de LTAS em frequência

alta e média (Tjaden et al., 2010; Ziemssen; Reichmann, 2010). Por outro lado, em idosos saudáveis, os parâmetros de qualidade de voz podem ser afetados devido ao arqueamento, atrofia das pregas vocais e flacidez na musculatura da laringe. A inclinação de LTAS em frequência média foi significativa para ambos os sexos. No grupo feminino, houve diferenças entre os estágios de severidade e entre os estilos de fala no terceiro estágio da DP. Os valores médios femininos são maiores nos subgrupos DP3 em ambos os estilos, seguidos pelos subgrupos DP2. Os homens somente se diferenciaram entre os estágios de severidade da DP. O terceiro estágio da DP teve os maiores valores médios, enquanto menores foram no segundo estágio da DP em ambos os estilos de fala.

O parâmetro de coeficiente de variação de intensidade, na DP, pode ser afetado devido ao sintoma de rigidez muscular que afeta a mobilidade dos músculos da laringe e devido à perda de amplitude de movimentos, o que pode influenciar na intensidade da voz. A postura curvada para a frente, que é comum em pessoas com DP, pode limitar a expansão do peito e reduzir a capacidade pulmonar (Ziemssen; Reichmann, 2010). Nos idosos saudáveis, por outro lado, esse parâmetro pode ser influenciado pela perda de massa muscular e pelo arqueamento das pregas vocais (Kost; Sataloff, 2018). Entre as mulheres, os subgrupos dos estágios da DP apresentaram valores mais elevados em relação ao grupo controle. As mulheres se diferenciaram somente entre os estágios da DP, e o terceiro estágio teve o maior valor médio do parâmetro.

No desvio padrão da primeira derivada positiva da frequência fundamental ( $df0sdpos$ ), observamos diferenças significativas somente nas mulheres, tanto para os estágios de severidade da DP quanto para os estilos de fala no terceiro estágio da DP. O subgrupo DP3-leitura teve o menor valor do parâmetro.

Na DP, os parâmetros temporais, como a taxa de elocução e a taxa de articulação, podem ser afetados devido à rigidez muscular, lentidão dos movimentos e demanda de maior esforço vocal (*emph*) para falar. A rigidez muscular na DP afeta a amplitude de movimento dos articuladores, como lábios e língua, levando a uma redução no contraste fônico. Isso pode resultar em um aumento na taxa de elocução e articulação e em inconsistências na produção do acento lexical primário à medida que a doença progride (Soares, 2009). Por outro lado, em idosos saudáveis, as alterações nesses parâmetros podem ser devido a uma combinação de fatores relacionados ao processo natural de envelhecimento que afeta a saúde física (e.g. alterações no controle motor e na força muscular) e a cognição (e.g. declínio cognitivo) (Manderson et al., 2022). A análise estatística revelou diferenças significativas na taxa de articulação no grupo feminino, com uma redução progressiva do valor médio de acordo com o grau de severidade da DP no estilo leitura. Entre os homens, as diferenças foram devidas aos estágios de severidade da DP para a taxa de articulação.

As medidas da frequência fundamental da fala podem ser afetadas pelos sintomas da DP, como a hipofonia, resultando em uma produção vocal reduzida. Variações na F0 podem causar uma voz monótona, enquanto o quadro de disartria pode levar a dificuldades no controle muscular para a produção da fala, afetando o ritmo e a modulação da frequência. Tremores podem causar ruídos na voz e contribuem com alterações na qualidade vocal (Rosa; Cielo; Cechella, 2009). Na mediana da frequência fundamental, o grupo masculino teve diferenças significativas tanto para os estágios de severidade da DP quanto para o grupo controle nos dois estilos de fala. O subgrupo DP3 nos dois estilos de fala apresentou os maiores valores

médios desse parâmetro, embora não tenha se diferenciado nos estilos de fala. No mínimo da frequência fundamental, os homens tiveram diferenças entre os estágios de severidade e entre os subgrupos controles para os estilos de fala. O subgrupo masculino DP3-comentário apresentou o maior valor médio.

O parâmetro de qualidade vocal *shimmer* pode estar alterado devido à variação na amplitude ciclo-a-ciclo e à irregularidade do contato das pregas vocais causados pelos sintomas de rigidez e bradicinesia (Rosa; Cielo; Cechella, 2009). Por outro lado, entre os idosos saudáveis, esse parâmetro é afetado devido ao arqueamento das pregas vocais e à flacidez da musculatura da laringe (Gomes; Simões-Zenari; Nembr, 2021). Os homens se diferenciaram tanto entre os estágios de severidade quanto entre os estilos de fala, sendo o subgrupo do estágio mais avançado os com maiores valores médios nos dois estilos.

Diferentemente de outros estudos (Lowit et al., 2018), não encontramos diferenças significativas entre as pausas. Por outro lado, as alterações em taxa de articulação, parâmetros da F0 e intensidade corroboram com os estudos (Patel; Parveen; Anand, 2016) que analisaram alterações de parâmetros prosódico-acústicos na DP.

## 6 CONCLUSÃO

O objetivo deste tudo foi analisar as diferenças nos parâmetros acústicos de diferentes dimensões fonéticas em discursos espontâneos e leitura em voz alta entre idosos. À luz da análise acústica, os resultados revelaram que certos parâmetros prosódico-acústicos, como a F0, ênfase espectral, taxas de articulação e elocução e *shimmer* podem ser afetados devido às restrições motoras da DP.

As limitações incluem a amostra geográfica única, o que restringe a generalização para populações diversas. Além disso, a natureza transversal do estudo e as condições específicas de coleta de dados impedem inferências sobre mudanças ao longo do tempo. Isso destaca a necessidade de estudos longitudinais com amostras mais diversas para entender melhor os efeitos da DP e do envelhecimento vocal.

## REFERÊNCIAS

- Ackermann H, Konczak J, Hertrich I. The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and language*. 1997;56(2):312-319.
- Appakaya SB, Sankar R. Classification of Parkinson's disease using pitch synchronous speech analysis. 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Jul. 2018; Honolulu, HI, USA: EMBC; 2018, p. 1420-1423.
- Armstrong MJ, Okun MS. Diagnosis and Treatment of Parkinson Disease. *JAMA*. 2020;323(6):548-560.
- Barbosa MT, et al. Parkinsonism and Parkinson's disease in the elderly: a community-based survey in Brazil (the Bambuí study). *Movement Disorders*. 2006;21(6):800-808.
- Barbosa P. Prosody Descriptor Extractor 2023 [internet].
- Barbosa P. Revelar a estrutura rítmica de uma língua construindo máquinas falantes: pela integração de ciência e tecnologia de fala. In: Scarpa EM, organizadora. *Estudos de prosódia*. Campinas: Editora da Unicamp; 1999. p. 21-52.

- Braak H, et al. Staging of the intracerebral inclusion body pathology associated with idiopathic Parkinson's disease (preclinical and clinical stages). *Journal of Neurology*. 2002;249(3 Suppl):iii1-iii5.
- Boersma, P; Weenink, D. Praat: doing phonetics by computer [programa de computador]. Versão 6.4.07.
- Cascudo LC. História da alimentação no Brasil. 3.<sup>a</sup> ed. São Paulo: Global; 2004.
- Dao SVT, et al. An analysis of vocal features for Parkinson's disease classification using evolutionary algorithms. *Diagnostics*. 2022;12(8):1980.
- Dromey C. Spectral measures and perceptual ratings of hypokinetic dysarthria. *Journal of Medical Speech-Language Pathology*. 2003;11:85-94.
- Dytham C. Choosing and using statistics: a biologist's guide. 3.<sup>a</sup> ed. Chichester: John Wiley & Sons; 2011.
- Fernandes I, Andrade Filho AS. Estudo clínico-epidemiológico de pacientes com doença de Parkinson em Salvador-Bahia. *Revista Brasileira de Neurologia e Psiquiatria*. 2018;22(1):45-59.
- Gomes ABP, Simões-Zenari M, Nemr K. Voz do idoso: o avanço da idade gera diferentes impactos? *CoDAS*. 2021;33(6): e20200126.
- Greffard S, et al. Motor score of the unified Parkinson disease rating scale as a good predictor of Lewy body-associated neuronal loss in the substantia nigra. *Archives of neurology*. 2006;63(4):584-588.
- Grosjean F, Lane H. Pauses and syntax in American Sign Language. *Cognition*. 1977;5:101-117.
- Hoehn MM, Yahr MD. Parkinsonism: onset, progression, and mortality. *Neurology*. 1967;17(5):427-442.
- Ikuta YM, et al. Avaliação da função cognitiva em pacientes com doença de Parkinson. *Rev. para med*. 2012;26(1):1-6.
- Kost KM, Sataloff RT. Voice disorders in the elderly. *Clinics in Geriatric Medicine*. 2018;34(2):191-203.
- Ladd DR. O que é prosódia? Couto TL, Seara IC, tradutoras. *Working Papers em Linguística*. 2019;20(1):8-46.
- Lirani-Silva C, Mourão LF, Gobbi LTB. Dysarthria and quality of life in neurologically healthy elderly and patients with Parkinson's disease. *CoDAS*. 2015;27(3):248-254.
- Lowit A, et al. Rhythmic performance in hypokinetic dysarthria: relationship between reading, spontaneous speech and diadochokinetic tasks. *Journal of Communication Disorders*. 2018;72:26-39.
- Ludlow CL, Connor NP, Bassich CJ. Speech timing in Parkinson's and Huntington's disease. *Brain and Language*. 1987;32(2):195-214.
- Manderson L, et al. A systematic review of the relationships amongst older adults' cognitive and motor speech abilities [internet].
- Maryn Y, et al. Mobile communication devices, ambient noise, and acoustic voice measures. *Journal of Voice*. 2017;31(2):248.e11-248.e23.
- Monnin P, Grosjean F. Les structures de performance en Français: caractérisation et prédiction. *L'année psychologique*. 1993;93(1):9-30.
- Pagano M, Gauvreau K. Principles of biostatistics. 2.<sup>a</sup> ed. New York: Chapman and Hall/CRC; 2018.
- Patel S, Parveen S, Anand S. Prosodic changes in Parkinson's disease. *Journal of the Acoustical Society of America*. 2016;140(4 Suppl):3442-3442.
- Quedas A, Duprat AC, Gasparini G. Lombard's effect's implication in intensity, fundamental frequency and stability on the voice of individuals with Parkinson's disease. *Brazilian Journal of Otorhinolaryngology*. 2007;73(5):675-683.
- RStudio Team. RStudio: integrated development for R [programa de computador]. Boston, MA: RStudio, PBC; 2023.

Rosa JDC, Cielo CAC, Cechella C. Função fonatória em pacientes com doença de Parkinson: uso de instrumento de sopro. *Revista CEFAC*. 2009;11(2):305–313.

Soares MFP. Estratégias de produção de vogais e fricativas: análise acústica da fala de sujeitos portadores de doença de Parkinson [tese]. Campinas: Instituto de Estudos da Linguagem, Universidade Estadual de Campinas; 2009.

Tanner K, et al. Spectral moments of the long-term average spectrum: sensitive indices of voice change after therapy? *Journal of Voice*. 2005;19(2):211–222.

Tjaden K, et al. Long-term average spectral (LTAS) measures of Dysarthria and their relationship to perceived severity. *Journal of Medical-Speech Language Pathology*. 2010;18(4):125-132.

Traunmüller H, Eriksson A. Acoustic effects of variation in vocal effort by men, women, and children. *The Journal of the Acoustical Society of America*. 2000;107(6):3438-3451.

Vieira MAS. Acento tonal pré-nuclear ascendente no português brasileiro: comparação com fala parkinsoniana [dissertação]. Belo Horizonte: Faculdade de Letras, Universidade Federal de Minas Gerais; 2017.

WEF Máquinas de Embalagens. Saiba como surgiu a paçoca - doce tradicionalmente brasileiro [internet].

Zardeto-Sabec G, et al. Doença de Parkinson, seus mecanismos fisiopatológicos e semiologia: revisão bibliográfica. *Brazilian Journal of Surgery and Clinical Research - BJSCR*. 2018;22(3):68-75.

Ziemssen T, Reichmann H. Treatment of dysautonomia in extrapyramidal disorders. *Therapeutic Advances in Neurological Disorders*. 2010;3(1):53-67.