

Marcia dos Santos Machado Vieira  
Marcos Luiz Wiedemer  
organizadores

# DIMENSÕES E EXPERIÊNCIAS EM SOCIOLINGUÍSTICA



MARCIA DOS SANTOS MACHADO VIEIRA  
MARCOS LUIZ WIEDEMER

(Organizadores)

# DIMENSÕES E EXPERIÊNCIAS EM SOCIOLINGUÍSTICA



2019

**Blucher**

 **CAPES**

**Blucher**

MARCIA DOS SANTOS MACHADO VIEIRA  
MARCOS LUIZ WIEDEMER  
(Organizadores)

# DIMENSÕES E EXPERIÊNCIAS EM SOCIOLINGUÍSTICA

*Albert Rilliard • Alessandra de Paula • Catherine Lee • Christina Abreu  
Gomes • Cláudia Regina Brescancini • Clinton Kakela Awai • Danielle  
Kely Gomes • Dinah Callou • Jonny Kim • Juliana Bertucci Barbosa •  
Katie Drager • Livia Oushiro • M. Joelle Kirtley • Marcia dos Santos  
Machado Vieira • Marcelo Alexandre Silva Lopes de Melo • Marcos Luiz  
Wiedemer • Silvia Figueiredo Brandão • Talita de Cássia Marine • Valéria  
Neto de Oliveira Monaretto*

*Dimensões e experiências em Sociolinguística*

© 2019 Marcia dos Santos Machado Vieira, Marcos Luiz Wiedemer  
Editora Edgard Blücher Ltda.

O presente livro foi produzido com apoio da Coordenação de Aperfeiçoamento de  
Pessoa de Nível Superior – Brasil (CAPES). Código Financiamento 001.

Diagramação: Laércio Flenic Fernandes

Revisão: Davi Pacheco Alves de Souza

---

# Blucher

Rua Pedroso Alvarenga, 1245, 4º andar  
04531-934 – São Paulo – SP – Brasil  
Tel 55 11 3078-5366  
contato@blucher.com.br  
www.blucher.com.br

Segundo Novo Acordo Ortográfico, conforme 5. ed.  
do Vocabulário Ortográfico da Língua Portuguesa,  
Academia Brasileira de Letras, março de 2009.

É proibida a reprodução total ou parcial por quaisquer  
meios, sem autorização escrita da Editora.

---

Todos os direitos reservados pela Editora  
Edgard Blücher Ltda.

Dados Internacionais de Catalogação na Publicação (CIP)  
Angélica Ilacqua CRB-8/7057

---

Dimensões e experiências em Sociolinguística  
Marcia dos Santos Machado Vieira, Marcos Luiz  
Wiedemer (orgs.). – São Paulo: Blucher, 2019.-  
314p.

Bibliografia

ISBN 978-85-212-1874-6 (e-book)  
ISBN 978-85-212-1875-3 (impresso)

Open Access

1. Sociolinguística I. Machado Vieira, Marcia dos  
Santos II. Wiedemer, Marcos Luiz

---

19-2023

CDD410

Índices para catálogo sistemático:  
1. Sociolinguística

---

## COMITÊ CIENTÍFICO AVALIADOR

Augusto Soares da Silva (Universidade Católica Portuguesa/Braga)  
Carla Valeria de Souza Faria (Università Ca' Foscari Venezia)  
Denise Cristina Kluge (Universidade Federal do Rio de Janeiro)  
Dermeval da Hora Oliveira (Universidade Federal da Paraíba / Presidente da Associação de Linguística e Filologia da América Latina)  
Heliana Ribeiro de Mello (Universidade Federal de Minas Gerais)  
Maria Jussara Abraçado de Almeida (Universidade Federal Fluminense)  
Laura Alvarez Lopez (Stockholm University)  
Marcio Martins Leitão (Universidade Federal da Paraíba)  
Maria Celeste Matias Rodrigues (Universidade de Lisboa)  
Maria da Conceição Paiva (Universidade Federal do Rio de Janeiro)  
Roberval Teixeira e Silva (Universidade de Macau)  
Vanessa Meireles de Oliveira Silva (Université Paul Valéry/Montpellier 3)

Proibida a reprodução total ou parcial em qualquer mídia sem autorização escrita da editora.  
Os infratores estão sujeitos às penas da lei.



*“Existem muitas formas de olhar para o mundo. Melhor do que chegar a um pressuposto fim onde tudo é um, é celebrar a pluralidade do saber, a natureza instável do conhecimento, fonte de nosso desejo de querer sempre buscar. Aceitar a incompletude do saber não é uma atitude derrotista; pelo contrário, é libertadora, pois entende que a busca não tem fim. E o que pode ser mais instigante do que saber que sempre existirá algo novo a ser descoberto?”.*

*(Marcelo Gleiser, 2019)<sup>1</sup>*

---

<sup>1</sup> GLEISER, Marcelo. *O caldeirão azul: o universo, o homem e seu espírito*. Rio de Janeiro: Record, 2019.



---

# SUMÁRIO

<b>PREFÁCIO/PREFACE.....</b>	<b>13</b>
<i>Cláudia Regina Brescancini (PUC-RS)</i>	
<i>Valéria Neto de Oliveira Monaretto (UFRGS)</i>	
<b>APRESENTAÇÃO/PRESENTATION.....</b>	<b>17</b>
<i>Marcia dos Santos Machado Vieira (UFRJ)</i>	
<i>Marcos Luiz Wiedemer (UERJ)</i>	
<b>1. SPEAKER ETHNICITY, LANGUAGE BACKGROUND, AND THE PRONUNCIATION OF HAWAIIAN PLACE NAMES.....</b>	<b>33</b>
<i>Katie Drager (University of Hawai‘i at Mānoa)</i>	
<i>M. Joelle Kirtley (University of Hawai‘i at Mānoa)</i>	
<i>Clinton Kakela Awai (The Ohio State University)</i>	
<i>Catherine Lee (University of Hawai‘i at Mānoa)</i>	
<i>Jonny Kim (Hanyang University)</i>	
<b>2. GEOPROSDY: QUANTITATIVE APPROACHES OF PROSODIC VARIATION ACROSS DIALECTS.....</b>	<b>55</b>
<i>Albert Rilliard (LIMSI, CNRS, Université Paris Saclay; UFRJ)</i>	
<b>3. SOCIOLINGUÍSTICA VARIACIONISTA E GRAMÁTICA DE CONSTRUÇÕES: OS DESAFIOS E AS PERSPECTIVAS DE COMPATIBILIZAÇÃO.....</b>	<b>85</b>
<i>Marcia dos Santos Machado Vieira (UFRJ)</i>	
<i>Marcos Luiz Wiedemer (UERJ)</i>	
<b>4. VARIATIONIST SOCIOLINGUISTICS AND CONSTRUCTION GRAMMAR: THE CHALLENGES AND THE PROSPECTS OF COMPATIBILIZATION.....</b>	<b>121</b>
<i>Marcia dos Santos Machado Vieira (UFRJ)</i>	
<i>Marcos Luiz Wiedemer (UERJ)</i>	
<b>5. PERCEPÇÃO DA VARIAÇÃO DA CODA (S) NA COMUNIDADE DE FALA DO RIO DE JANEIRO: ACESSANDO O SIGNIFICADO SOCIAL DA VARIANTE FRICATIVA POSTERIOR.....</b>	<b>129</b>
<i>Marcelo Alexandre Silva Lopes de Melo (UFRJ)</i>	
<i>Christina Abreu Gomes (UFRJ/CNPq)</i>	

<b>6. PERCEPTION OF THE VARIATION OF THE CODA (S) IN THE SPEECH COMMUNITY OF RIO DE JANEIRO: ACCESSING THE SOCIAL MEANING OF THE BACK-FRICATIVE VARIANT.....</b>	<b>149</b>
<i>Marcelo Alexandre Silva Lopes de Melo (UFRJ)</i> <i>Christina Abreu Gomes (UFRJ/CNPq)</i>	
<b>7. QUESTÕES E MÉTODOS: VOGAIS MÉDIAS PRETÔNICAS NA FALA DE MIGRANTES NORDESTINOS EM SITUAÇÃO DE CONTATO DIALETAL.....</b>	<b>157</b>
<i>Livia Oushiro (Unicamp)</i>	
<b>8. QUESTIONS AND METHODS: PRETONIC MIDVOWELS IN THE SPEECH OF NORTHEASTERN MIGRANTS IN DIALECTAL CONTACT .....</b>	<b>185</b>
<i>Livia Oushiro (Unicamp)</i>	
<b>9. O APAGAMENTO DA VOGAL POSTÔNICA MEDIAL EM DUAS VARIEDADES AFRICANAS DO PORTUGUÊS: UMA DESCRIÇÃO PRELIMINAR .....</b>	<b>189</b>
<i>Danielle Kely Gomes (UFRJ)</i>	
<b>10. THE DELETION OF POSTONIC MEDIAL VOWEL IN TWO AFRICAN VARIETIES OF PORTUGUESE: A PRELIMINARY STUDY .....</b>	<b>207</b>
<i>Danielle Kely Gomes (UFRJ)</i>	
<b>11. AINDA SOBRE OS RÓTICOS NO PORTUGUÊS DE MOÇAMBIQUE.....</b>	<b>211</b>
<i>Silvia Figueiredo Brandão (UFRJ/CNPq)</i> <i>Alessandra de Paula (UERJ/FAPERJ)</i>	
<b>12. STILL ON ROTHICS IN PORTUGUESE IN MOZAMBIQUE.....</b>	<b>237</b>
<i>Silvia Figueiredo Brandão (UFRJ/CNPq)</i> <i>Alessandra de Paula (UERJ/FAPERJ)</i>	
<b>13. PARA UMA NORMA-PADRÃO FLEXÍVEL NO CONTEXTO ESCOLAR: CONTRIBUIÇÕES DOS ESTUDOS SOCIOLINGUÍSTICOS .....</b>	<b>243</b>
<i>Silvia Rodrigues Vieira (UFRJ)</i>	
<b>14. TOWARDS A FLEXIBLE STANDARD IN THE SCHOOL CONTEXT: CONTRIBUTIONS FROM LINGUISTIC STUDIES.....</b>	<b>265</b>
<i>Silvia Rodrigues Vieira (UFRJ)</i>	
<b>15. ATITUDES LINGUÍSTICAS DE MESTRANDOS DAS UNIDADES DO PROFLETRAS DA REGIÃO DO TRIÂNGULO MINEIRO .....</b>	<b>269</b>
<i>Juliana Bertucci Barbosa (UFTM-Uberaba)</i> <i>Talita de Cássia Marine (UFU-Uberlândia)</i>	

<b>16. LINGUISTIC ATTITUDES OF MASTER STUDENTS OF PROFLETRAS UNITS IN THE "TRIÂNGULO MINEIRO" REGION .....</b>	<b>295</b>
<i>Juliana Bertucci Barbosa (UFTM-Uberaba)</i>	
<i>Talita de Cássia Marine (UFU-Uberlândia)</i>	
<b>ENTREVISTA COM A SOCIOLINGUISTA</b>	
<b>DINAH MARIA ISENSEE CALLOU .....</b>	<b>299</b>
<i>Marcos Luiz Wiedemer (UERJ)</i>	
<i>Marcia dos Santos Machado Vieira (UFRJ)</i>	
<b>OS ORGANIZADORES .....</b>	<b>303</b>
<b>OS AUTORES .....</b>	<b>305</b>



---

# PREFÁCIO

Esta obra reúne resultados de pesquisas desenvolvidas sobre temas de interesse de linguistas vinculados ao Grupo de Trabalho (GT) de Sociolinguística da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística (ANPOLL) e de outros linguistas convidados. Em seus quase 35 anos de existência, o GT, constituído por pesquisadores/docentes atuantes em Programas de Pós-Graduação em Letras e em Linguística de universidades brasileiras e por vários de seus orientandos, tem fomentado o desenvolvimento da Sociolinguística no país a partir do interesse nas diversas faces do estudo da língua em seu contexto social.

Os trabalhos aqui apresentados são representativos dos eixos temáticos que, desde 2014, constituem as principais linhas de investigação do GT. Esses eixos temáticos, Variação e Mudança Linguística; Contato, Variação e Identidade; Sociolinguística e Ensino e Teorias e Métodos para o Estudo da Variação e Mudança Linguísticas: estado da arte e perspectivas, agregaram diferentes perspectivas descritivas e teóricas atualmente adotadas pelos pesquisadores do grupo.

*Variação e Mudança Linguística*, inicialmente coordenado por Dinah Maria Isensee Callou (UFRJ), linguista entrevistada nesta obra, Maria Eugênia Lamoglia Duarte (UFRJ), Silvia Figueiredo Brandão (UFRJ) e Marcia Machado Vieira (UFRJ), e posteriormente, a partir de 2016, por Jacyra Andrade Mota (UFBA), Izete Lehmkuhl Coelho (UFSC), Silvia Figueiredo Brandão (UFRJ) e Marcia Machado Vieira (UFRJ), volta-se à descrição e análise de fenômenos no âmbito da fonética, fonologia e prosódia; morfologia e sintaxe; léxico; semântica, discurso e pragmática, e que versem sobre discussões (a) teóricas que permitam avaliar o alcance de diferentes teorias na aplicação do modelo variacionista para o estudo da mudança, particularmente no que se refere às restrições, implementação, encaixamento e transição; (b) metodológicas, com enfoque na avaliação e encaixamento social de variantes, na investigação sobre a origem das mudanças (de cima para baixo ou de baixo para cima) e na sua caracterização, como estáveis ou em progresso, e (c) sobre comunidades de práticas.

*Contato, Variação e Identidade*, coordenado por Mônica Maria Guimarães Savedra (UFF) e Luís Isaias Centeno do Amaral (UFPel) e, a partir de 2016, por Mônica Maria Guimarães Savedra (UFF) e Karen Spinassé (UFRGS), focaliza as relações entre variação linguística, contato linguístico, cultura e identidade. Envolve estudos de variação e de contato linguístico que tratem de dimensões socioculturais como gênero, faixa etária, etnia, classe, entre outros e que englobem a temática dos contatos entre o Português Brasileiro e línguas autóctones, alóctones, provenientes de contextos de colonização, imigração e de fronteiras linguísticas e geográficas. São também de interesse a esse eixo as discussões sobre avaliação, orientação e atitudes linguísticas.

*Sociolinguística e Ensino*, coordenado por Silvia Rodrigues Vieira (UFRJ) e Lúcia Cyranka (UFJF) e, a partir de 2018, por Silvia Rodrigues Vieira (UFRJ) e Joyce Baronas (UEL), visa à articulação entre os resultados dos estudos variacionistas e o contexto do ensino de Língua Portuguesa, a partir de uma agenda de trabalho que contempla (a) a sistematização do comportamento de variantes linguísticas quanto à modalidade (oral x escrita), à monitoração estilística, a variedades no espaço físico e social; (b) a investigação de crenças e atitudes em relação ao uso de variantes em cada fenômeno variável e (c) ao desenvolvimento de metodologias para o tratamento da variação em sala de aula.

*Teorias e Métodos para o Estudo da Variação e Mudança Linguísticas: estado da arte e perspectivas*, coordenado inicialmente por Marco Antonio Martins (UFSC) e Rosane Berlinck (UNESP) e, a partir de 2017, também por Maria Jussara Abraçado de Almeida (UFF), aborda a sistematização das contribuições evidenciadas nos muitos anos de estudo da variação linguística no Brasil e a direção de desenvolvimento do campo com base em questões que envolvem (a) os limites e ganhos das articulações teóricas possíveis, (b) o instrumental metodológico adotado e (c) a ampliação das investigações para novas fronteiras teóricas e metodológicas.

A relevância das pesquisas desenvolvidas no escopo dos eixos temáticos para a descrição e análise do Português Brasileiro, como se pode verificar nesta coletânea de artigos, reflete muito das frutíferas discussões conduzidas nas reuniões anuais dos membros do GT de Sociolinguística e, sobretudo, do amadurecimento desses pesquisadores como *grupo de trabalho*.

Cláudia Regina Brescancini (PUC-RS)  
Valéria Neto de Oliveira Monaretto (UFRGS)  
(Coordenadoras do GT de Sociolinguística – 2014-2018)

---

# PREFACE

This study gathers together results of research developed on topics of interest to linguists involved in the Working Group (WG) of Sociolinguistics from the National Association of Postgraduate Courses and Research in Letters and Linguistics (ANPOLL), and guests. In its almost 35 years of existence, the WG, consisting of researchers / professors working in graduate programs in Letters and Linguistics at Brazilian universities and several of their students, has characterized the development of Sociolinguistics in the country based on the interest in the several facets of language study in its social context.

The studies presented here are representative of the thematic axes which, from 2014, constitute the WG's main research lines. These thematic axes, Linguistic Variation and Change; Contact, Variation and Identity; Sociolinguistics and Teaching; and Theories and Methods for the Study of Linguistic Variation and Change: state of the art and perspectives, combine different descriptive and theoretical perspectives currently adopted by the WG's researchers.

*Linguistic Variation and Change*, initially coordinated by Dinah Maria Isensee Callou (UFRJ), linguist interviewed in this study, Maria Eugênia Lamoglia Duarte (UFRJ), Silvia Figueiredo Brandão (UFRJ) and Marcia Machado Vieira (UFRJ), and, from 2016, by Jacyra Andrade Mota (UFBA), Izete Lehmkuhl Coelho (UFSC), Silvia Figueiredo Brandão (UFRJ) and Marcia Machado Vieira (UFRJ). This thematic axis explores description and analysis of phenomena in the field of phonetics, phonology and prosody; morphology and syntax; lexicon; semantics, discourse and pragmatics which are about (a) theoretical discussions that allow the evaluation of the scope of different theories in applying the variationist model to the study of change, particularly regarding constraints, implementation, fitting and transition; (b) methodological discussions that focus on the evaluation and social fitting of variants in the research on the origin of changes (top-down or bottom-up) and their characterization as stable or in progress, and (c) discussions about communities of practices.

*Contact, Variation and Identity*, coordinated by Mônica Maria Guimarães Savedra (UFF) and Luís Isaias Centeno do Amaral (UFPel), and, from 2016, by

Mônica Maria Guimarães Savedra (UFF) and Karen Spinassé (UFRG). This thematic axis focuses on the relationships between linguistic variation, linguistic contact, culture and identity. It involves studies of variation and linguistic contact which deal with sociocultural dimensions such as gender, age, ethnicity, class, among others and which encompass the theme of contacts between Brazilian Portuguese and autochthonous languages, allochthonous languages, stemming from contexts of colonization, immigration and of linguistic and geographical boundaries. Discussions on assessment, orientation and linguistic attitudes are also of interest to this axis.

*Sociolinguistics and Teaching*, coordinated by Silvia Rodrigues Vieira (UFRJ) and Lúcia Cyranka (UFJF), and, from 2018, by Silvia Rodrigues Vieira (UFRJ) and Joyce Baronas (UEL). This thematic axis aims to connect the results of the variationist studies and the context of teaching of the Portuguese Language, from a work agenda which takes into consideration (a) systematization of linguistic behavior variants regarding the modality (oral x written), stylistic monitoring, varieties in the physical and social space; (b) research on beliefs and attitudes regarding the use of variants in each variable phenomenon; and (c) the development of methodologies for the treatment of classroom variation.

*Theories and Methods for the Study of Linguistic Variation and Change: state of the art and perspectives*, initially coordinated by Marco Antonio Martins (UFSC) and Rosane Berlinck (UNESP), and from 2017, also by Maria Jussara Abraçado de Almeida (UFF). This thematic axis addresses the systematization of the contributions evidenced in the many years of study of linguistic variation in Brazil and the direction of field development from issues which involve (a) the limits and gains of possible theoretical articulations, (b) the methodological tools adopted and (c) the expansion of investigations to new theoretical and methodological frontiers.

The relevance of research carried out within the scope of the thematic axes for the description and analysis of Brazilian Portuguese, as can be seen in this collection of articles, reflects much of the fruitful discussions conducted by members of the Sociolinguistics WG at the annual meetings and, especially, the maturation of these researchers as a *working group*.

*Claudia Regina Brescancini (PUC-RS)*  
*Valéria Neto de Oliveira Monaretto (UFGRS)*  
*(Coordinators of the Sociolinguistics WG – 2014-2018)*

---

# APRESENTAÇÃO

*Marcia dos Santos Machado Vieira(UFRJ)*  
*Marcos Luiz Wiedemer(UERJ)*

Esta obra congrega duas intenções de interlocução: por um lado, compartilhar com leitores no Brasil e fora do Brasil (pesquisadores ou leigos) um pouco das inquietações atuais e das conquistas da Sociolinguística Brasileira; por outro lado, divulgar dimensões e experiências na área de Sociolinguística, com o intuito de fazer prosperarem (i) pontes entre pesquisadores envolvidos em experiências diversas (nacional e internacionalmente) e (ii) diálogos sobre temáticas que interessam à investigação linguística no cenário nacional. Para tanto, consolida-se com base em sete capítulos escritos por pesquisadores sociolinguistas brasileiros, dois capítulos escritos por pesquisadores estrangeiros (da Universidade de Hawai'i e de LIMSI, CNRS, Université Paris Saclay; UFRJ) e uma entrevista feita pelos organizadores a uma docente-pesquisadora sociolinguista brasileira (da Universidade Federal do Rio Janeiro), expoente na área.

Este livro reúne pesquisas que foram apresentadas durante o encontro do Grupo de Trabalho (GT) de Sociolinguística da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística (ANPLL) – ocorrido nos dias 27, 28 e 29 de junho de 2018, na Universidade Federal do Mato Grosso (UFMT), na cidade de Cuiabá. O GT possui quatro eixos temáticos de trabalho e investigação, a saber: (i) *Variação e Mudança Linguística*; (ii) *Contato*,

*Varição e Identidade*; (iii) *Sociolinguística e Ensino* e (iv) *Questões Teóricas e Metodológicas*. Tais eixos temáticos têm por essência a diversidade de experiências de configuração do objeto observacional a partir do olhar sociolinguístico – articulado ou não a outro olhar teórico-metodológico. Desse modo, é inerente às pesquisas do GT assentarem-se em diferentes dimensões em potencial de observação num objeto linguístico e em diversas experiências para o mapear empiricamente ou para o apreender teórica e/ou metodologicamente. Pode-se ter uma ideia da riqueza de possibilidades de seu fazer científico por meio da representação das quatro frentes de trabalho do GT, a qual se faz presente nos capítulos desta obra.

Desde que a pesquisa linguística passou a observar o uso da língua dentro da comunidade de fala, levando em conta o contexto social de produção, a Sociolinguística propicia aos pesquisadores o desenvolvimento de diversas agendas de trabalho: desde agendas relacionadas a diferentes perspectivas de enfrentamento de questões linguísticas (ECKERT, 2012 e 2018)<sup>1</sup> na Teoria da Variação e Mudança Linguística, até agendas de investigação na Dialectologia/Geolinguística, na Sociolinguística Etnográfica, na Etnografia da Comunicação, na Sociologia da Linguagem, entre outras potencialidades. Destaque relevante nessa área diz respeito às interfaces teóricas que ela ensina, como Geossociolinguística, Sociolinguística Paramétrica, Sociofuncionalismo, e, mais recentemente, a Sociofonética, Sociolinguística Cognitiva e o Socioconstrucionismo. Outro destaque na área relaciona-se aos trabalhos que discutem as questões de Contato linguístico, línguas minoritárias e representação de identidade, bem como as diversas contribuições e diálogos com os campos do Ensino e de Políticas Linguísticas. Assim, sem tencionarmos captar toda a potencialidade da área nesta breve apresentação e na obra, apresentamos aqui um pouco dos perfis da Sociolinguística brasileira e da capacidade do GT de manter conexão entre a atenção ao conhecimento já acumulado em razão do trabalho científico de várias gerações de sociolinguistas e a atenção ao que há de novo em termos de ideias, frentes de pesquisa, escopos teórico-metodológicos, técnicas de mapeamento ou tratamento de dados.

Sob a organização dos professores Marcia dos Santos Machado Vieira (Coordenadora) e Marcos Luiz Wiedemer (Vice-Coodenador), na gestão do GT de

---

<sup>1</sup> ECKERT, Penelope. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *The Annual Review of Anthropology*, 41:87–100, 2012. // ECKERT, Penelope. *Meaning and Linguistic Variation: The Third Wave in Sociolinguistics*. Cambridge University Press, 2018.

Sociolinguística (2018-2020), os capítulos aqui reunidos demonstram, justamente, as diferentes dimensões de pesquisas da área, bem como convidam o leitor a se engajar numa interlocução sociolinguística sobre a língua e experiências construídas a partir de questões por ela suscitadas. Assim, materializando-se como uma importante contribuição desse GT, a obra tem por objetivo aprofundar alguns temas de investigações em curso e, então, responder a algumas questões da área. O conjunto de pesquisas que compõem o livro traz três demandas urgentes para a reflexão a que se convida o leitor: (i) a natureza das relações sociais e de suas interações e a forma como são representadas nas línguas; (ii) o desenvolvimento de novos enquadres teóricos e a definição de diretrizes e técnicas para novas pesquisas; e (iii) a concepção e efetivação de estratégias político-pedagógicas que considerem o caráter heterogêneo da língua, bem como o multilinguismo de espaços (geo-)sociais. Tais demandas são, em alguma medida, debatidas nos capítulos que compõem a presente coletânea e que, então, passamos a apresentar resumidamente.

O primeiro capítulo deste livro, “*Speaker ethnicity, language background, and the pronunciation of Hawaiian place names*”, é de autoria dos pesquisadores Katie Drager, M. Joelle Kirtley, Clinton Kakela Awai, Catherine Lee e Jonny Kim e apresenta um estudo realizado com os moradores do Havaí, que exibem uma grande variação na pronúncia de nomes de lugares que têm origem havaiana, focalizando nestas variáveis linguísticas: a oclusiva glotal e a realização da vogal /o/. Os pesquisadores investigam se a realização fonética desses nomes de lugares havaianos é vinculada à etnia do falante ou à fala havaiana. Após a aplicação de rigorosa pesquisa empírica, os resultados do estudo demonstram evidências de que ambas as variáveis – tanto a origem étnica do falante quanto a sua capacidade de falar havaiano – estão relacionadas às variantes fonéticas usadas. Os autores argumentam que essas realizações fonéticas são indexadas a significados sociais de reivindicações dos havaianos. A pesquisa ressalta, ainda, que os falantes podem, através da produção da fala, reconhecer as origens indígenas desses nomes de lugares, bem como associar implicitamente os lugares ao povo e à cultura havaiana.

O segundo capítulo, com o título “*Geoprosody: quantitative approaches of prosodic variation across dialects*”, de autoria do pesquisador francês Albert Rilliard, propõe, através da abordagem denominada Geoprosódia, uma revisão dos métodos usados para medir as divergências objetivas entre os desempenhos prosódicos, tipicamente em relação à variação dialetológica. O ponto focal de observação está no fato de que as mudanças nas estruturas prosódicas dos dialetos

raramente são focalizadas nessas abordagens dialetométricas. Com isso, o autor discute, com foco nas línguas românicas, os métodos adequados para medir mudanças prosódicas entre variedades linguísticas, bem como a adequação das diferentes medidas a vários tipos de análises, sua robustez ao ruído de medição, sua aplicabilidade e os diferentes tipos de conjuntos de dados.

Esses dois capítulos atendem, em certa medida, a anseios manifestos por membros do GT de Sociolinguística no encontro de 2018 e relacionados a questões no âmbito das temáticas *avaliação subjetiva/percepção e dialectologia/geolinguística*.

O capítulo “*Sociolinguística Variacionista e Gramática de Construções: o desafio e as perspectivas de compatibilização*”, de Marcia dos Santos Machado Vieira e Marcos Luiz Wiedemer, parte do pressuposto de que é preciso pôr em foco o tratamento da variação no modelo da Gramática de Construções Baseada no Uso. Esse intento é consubstanciado pela articulação desse referencial com o da Sociolinguística Variacionista em razão da experiência alcançada por esta (conforme sugerem BYBEE, em entrevista a TORRENT, 2012<sup>2</sup>, e HOFFMAN; TROUSDALE, 2011<sup>3</sup>). Para fomentar a defesa dessa articulação, os autores concentram-se em um esforço para o refinamento do conceito de “variação construcional” e de como as abordagens construcionistas centradas em dados (usos ou registros de avaliação subjetiva/percepção de usos) podem lidar com esse fenômeno. O foco específico do capítulo é a problematização de duas dentre as possibilidades de variação construcional – a variação por analogia/alinhamento de construções independentes e a variação por compatibilização de (co)lexemas a *slot* de construção –, uma novidade em termos teórico-metodológicos no Brasil, que só atualmente começa a chamar a atenção de construcionistas funcionalistas brasileiros. E, para ilustrar a configuração dessas duas possibilidades de variação, os autores recorrem a construções de predicação verbal no Português.

“*Percepção da variação da coda (s) na comunidade de fala do Rio de Janeiro: acessando o significado social da variante fricativa posterior*”, capítulo de autoria de Marcelo Alexandre Silva Lopes de Melo e Christina Abreu Gomes, apresenta os resultados de um teste de percepção que avalia o

---

<sup>2</sup> TORRENT, Tiago Torrent. Entrevista. Usage-based models in linguistics: an interview with Joan Bybee. *Revista LinguiStica*, vol. 8, 1, p. 1-6, 2012.

<sup>3</sup> HOFFMANN, Thomas; TROUSDALE, Graeme. (Eds.). Special issue of Cognitive Linguistics: Variation, change and constructions in English. *Cognitive Linguistics*, vol. 22, n. 1, p. 1-23, 2011.

significado social das variantes da variável coda (s) aplicado a três grupos de participantes pertencentes a dois grupos sociais distintos na comunidade de fala do Rio Janeiro. A partir da premissa de que a variante fricativa posterior (velar/glotal) é estigmatizada em diversas variedades do português brasileiro, os resultados da pesquisa revelam que não só existem diferentes avaliações das variantes da coda (s) entre os diferentes grupos de participantes, como também as diferenças são graduais. O estudo demonstra, ainda, que padrões de avaliação estão intimamente relacionados ao grau de acesso dos indivíduos dos diferentes grupos às instituições formadoras dos padrões de prestígio.

Livia Oushiro, no capítulo “*Questões e métodos: vogais médias pretônicas na fala de migrantes nordestinos em situação de contato dialetal*”, analisa padrões de variação na altura das vogais médias pretônicas /e/ e /o/ na fala de migrantes paraibanos e alagoanos residentes na cidade de São Paulo e discute métodos para o tratamento de dados sociofonéticos. Os resultados mostram uma correlação entre a altura das vogais e o tempo de migração; e, com isso, considerando as análises a partir dos métodos empregados, a autora avalia que o campo de estudos variacionistas, no Brasil, necessita de uma mudança de postura quanto ao tratamento de dados sociolinguísticos.

O texto “*O apagamento da vogal postônica medial em duas variedades africanas do português: uma descrição preliminar*”, de Danielle Kely Gomes, apresenta resultados de análises do processo de apagamento da vogal postônica medial, o qual culmina na regularização de proparoxítonos ao padrão paroxítono. Procurando demonstrar novas evidências para o fenômeno em pauta, a autora apresenta uma comparação entre dados do Português de São Tomé e do Português de Moçambique, a partir da hipótese de que as proparoxítonas, não naturais até para falantes de Português como L1, seriam consistentemente regularizadas a paroxítonas nas duas variedades, como efeito do contato do Português com as outras línguas que com ele coexistem em ambas as comunidades.

Silvia Figueiredo Brandão e Alessandra de Paula, no capítulo “*Ainda sobre os róticos no Português de Moçambique*”, investigam quais fatores estruturais e sociais condicionam o uso das variantes de R na posição pré-vocálica, na fala de indivíduos que utilizam o Português como L1 e L2, na variedade urbana do Português de Moçambique. As autoras argumentam que tudo indica que, na gramática do Português de Moçambique, há um único fonema rótico, tendo em vista os resultados das análises demonstrarem que o tepe, a variante dominante nessa variedade, ocorre na fala de todos os indivíduos, embora seja significativamente mais frequente entre aqueles que o têm como L2. Entre os falantes de PM L1,

sobretudo os de nível superior de escolaridade, prevalece a vibrante alveolar, o que aproxima a norma deles da norma europeia.

O capítulo escrito por Silvia Rodrigues Vieira, que tem como título “*Para uma norma-padrão flexível no contexto escolar: contribuições dos estudos linguísticos*”, parte da premissa de que as normas praticadas em diversas variedades/modalidades precisam admitir natural flexibilidade e adaptabilidade por contexto sociocomunicativo. A autora procura aprofundar a discussão do estabelecimento da norma-padrão. Para tanto, apresenta orientações para o tratamento de trabalhos pedagógicos que considerem a variação. Além disso, é destacada a pertinência da relação entre orientação normativa e conhecimento sociolinguisticamente fundamentado.

Juliana Bertucci Barbosa e Talita de Cássia Marine, em “*Atitudes linguísticas de mestrandos das unidades do PROFLETRAS da região do Triângulo Mineiro*”, buscam tecer discussão respaldada por resultados de teste de atitude linguística aplicado a professores da Educação Básica da rede pública de ensino, ingressantes no Programa de Mestrado Profissional em Letras (PROFLETRAS), de duas unidades da região do Triângulo Mineiro. Para o teste, as autoras consideram dois fenômenos variáveis do Português Brasileiro: concordância verbal (mais estigmatizado) e alternância entre os verbos “ter” e “haver” com sentido de existir (menos estigmatizado). O intuito no capítulo é compreender sociolinguisticamente as atitudes linguísticas desses professores e subsidiar resultados para o planejamento de ações educacionais.

Cada um dos capítulos elaborados por sociolinguistas brasileiros é seguido de um resumo ampliado escrito em inglês. Esse cuidado na organização desta obra visa a contribuir para a circulação mais ampla possível de práticas de investigação e saberes sobre a língua portuguesa.

Os organizadores oferecem aos leitores, ainda, o texto “*Entrevista com a Sociolinguista Dinah Maria Isensee Callou*” (Professora Emérita da UFRJ) a partir de dez questões. Nela a pesquisadora trata da importância da área da Sociolinguística no âmbito das pesquisas linguísticas, das contribuições e repercussões dessa área nos cenários nacional e internacional, de seu caráter interdisciplinar, do valor do trabalho em grupo, bem como de outros temas.

Cientes da importância da nossa constante interlocução com interessados em investigações sociolinguísticas de dentro e de fora das áreas de Letras e Linguística, bem como de dentro e de fora do Brasil, os organizadores deste livro procuram fomentar, em linhas gerais, a relação *pesquisa de língua e sociedade*, por meio das escolhas feitas em prol da configuração desta obra e de iniciativas

como a de constituir um fórum<sup>4</sup> cujo intuito é viabilizar interlocuções que se desenvolvam sob os formatos mais diversos e que, então, atendam, na medida do possível, a demandas dos interessados.

Agradecemos imensamente aos autores, que se dispuseram a contribuir para a realização deste livro, e ao Comitê Científico Avaliador (composto por docentes pesquisadores brasileiros e estrangeiros, que tão prontamente atenderam ao nosso convite), pela interlocução tão produtiva e pelas contribuições para o importante “(re)fazer-se”/aprimoramento envolvido nesse processo, no nosso trabalho.

Agradecemos, especialmente, à Coordenação de Aperfeiçoamento de Pessoal Nível Superior (CAPES), pelo apoio concedido. Com esse apoio, acolheu uma proposta de visibilidade da ciência linguística brasileira que se produz e que tanto tem contribuído para o conhecimento nas áreas de Letras e Linguística, bem como para a formação de quadros acadêmico-profissionais, não só no país, que fazem diferença no dia a dia das comunidades em que esse conhecimento assume feições de prática(s). Agradecemos, ademais, ao Programa de Pós-Graduação em Letras Vernáculas da UFRJ e ao Programa de Pós-Graduação em Letras e Linguística da UERJ, por também acolherem essa ideia, bem como por apoiarem esta obra e, então, viabilizarem a parceria interinstitucional materializada no processo de produção deste livro.

Resta-nos, por fim, convidar todo e qualquer leitor interessado em questões sociolinguísticas a ler e compartilhar as dimensões e experiências trazidas nos capítulos desta obra, bem como a interagir com o GT de Sociolinguística da ANPOLL (<http://anpoll.org.br/>) e com os membros do Fórum Internacional em Sociolinguística ([www.fis2019.com.br](http://www.fis2019.com.br)).

Boa leitura!

*Os organizadores  
Rio de Janeiro (RJ), agosto de 2019.*

---

<sup>4</sup> Em 2019, o FIS (Fórum Internacional em Sociolinguística, [www.fis2019.com.br](http://www.fis2019.com.br)) oferece à comunidade um evento presencial (na Faculdade de Letras da UFRJ), concebido pela gestão do GT de Sociolinguística da ANPOLL, no biênio 2018-2020, que conta com o apoio da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), Edital PAEP, Projeto no. 8887.359599/2019-00, auxílio 865/2019); do Programa de Pós-Graduação em Letras Vernáculas da UFRJ, do Programa de Pós-Graduação em Letras e Linguística da UERJ, bem como com a parceria da Universidade Federal do Rio de Janeiro, Universidade do Estado do Rio de Janeiro, Università Ca' Foscari Venezia e Université Paul Valéry/Montpellier 3.



---

# PRESENTATION

*Marcia dos Santos Machado Vieira (UFRJ)*

*Marcos Luiz Wiedemer (UERJ)*

This book brings together two intentions of dialogue: on the one hand, to share with readers in Brazil and outside Brazil (researchers or laypeople) a little of the current concerns and achievements of Brazilian Sociolinguistics; on the other hand, to disseminate dimensions and experiences in the area of Sociolinguistics, with the purpose of promoting (i) bridges between researchers involved in diverse experiences (nationally and internationally) and (ii) dialogues on themes that are of interest to the linguistic research in the national scenario. To this end, it is consolidated on the basis of seven chapters written by Brazilian sociolinguistic researchers, two chapters written by foreign researchers (from the University of Hawai'i and the LIMSI, CNRS, Université Paris Saclay; UFRJ) and an interview conducted by the organizers with a Brazilian sociolinguistic professor-researcher (from the Federal University of Rio de Janeiro), an exponent in the area.

This book gathers researches that were presented during the meeting of the Sociolinguistics Working Group (WG) of the National Association of Postgraduate Courses and Research in Letters and Linguistics (ANPOLL) – held on June 27, 28 and 29, 2018, at the Federal University of Mato Grosso (UFMT), in Cuiabá city. The WG has four thematic axes of work and research, namely: (i) *Linguistic*

*Variation and Change*; (ii) *Contact, Variation and Identity*; (iii) *Sociolinguistics and Teaching*; and (iv) *Theoretical and Methodological Issues*. Such thematic axes have as their essence the diversity of observational object configuration experiences from the sociolinguistic view – articulated or not to another theoretical-methodological view. Thus, it is inherent in the GT researches to be based on different dimensions that can be potentially observed in a linguistic object and in various experiences by empirically mapping it or grasping it theoretically and/or methodologically. One can get an idea of the abundance of possibilities of its scientific work through the representation of the WG four work fronts which is embodied in the chapters of this book.

Since linguistic research has begun to observe the use of language within the speech community, taking into account the social context of production, Sociolinguistics has enabled researchers to develop diverse work agendas: from agendas related to different perspectives of coping with linguistic issues (ECKERT, 2012 and 2018)<sup>1</sup> in the Theory of Linguistic Variation and Change, up to research agendas in Dialectology/Geolinguistics, Ethnographic Sociolinguistics, The Ethnography of Communication, The Sociology of Language, among other potentialities. Relevant highlight in this area concerns the theoretical interfaces it entails, such as Geosociolinguistics, Parametric Sociolinguistics, Sociofunctionalism, and, more recently, Sociophonetic, Cognitive Sociolinguistics, and Socioconstructionism. Another highlight in the area is related to the works that discuss the issues of Linguistic Contact, minority languages and identity representation, as well as the various contributions and dialogues with the fields of Teaching and Language Policies. Thus, without intending to capture the full potential of the area in this brief presentation and in the work, we present here some of the profiles of the Brazilian Sociolinguistics and the WG's capacity to maintain a connection between the attention to the knowledge already accumulated due to the scientific work of several generations of sociolinguists and attention to what is new in terms of ideas, research fronts, theoretical-methodological scopes, mapping techniques or data processing.

Under the organization of professors Marcia dos Santos Machado Vieira (Coordinator) and Marcos Luiz Wiedemer (Vice-Coordinator), in the management of the Sociolinguistics WG (2018-2020), the chapters gathered here de-

---

<sup>1</sup> ECKERT, Penelope. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. *The Annual Review of Anthropology*, 41:87–100, 2012. // ECKERT, Penelope. *Meaning and Linguistic Variation: The Third Wave in Sociolinguistics*. Cambridge University Press, 2018.

monstrate precisely the different dimensions of research in the area and invite the reader to engage in a sociolinguistic dialogue about the language and experiences built on questions about it. Thus, materializing as an important contribution of this WG, the work aims to deepen some themes of ongoing investigations and then answer some questions in the area. The set of researches that makes up the book brings three urgent demands for the thinking to which the reader is invited: (i) the nature of the social relations and their interactions and the way they are represented in languages; (ii) the development of new theoretical frameworks and the definition of guidelines and techniques for further research; and (iii) the conception and implementation of political-pedagogical strategies that consider the heterogeneous character of the language, as well as the multilingualism of (geo)social spaces. Such demands are, to some extent, debated in the chapters that make up the present collection, which we summarize from now on.

The first chapter of this book, *“Speaker Ethnicity, Language Background, and the Pronunciation of Hawaiian Place Names,”* is authored by the researchers Katie Drager, M. Joelle Kirtley, Clinton Kakela Awai, Catherine Lee, and Jonny Kim (from an American university, University of Hawai‘i at Mānoa). It presents a study about Hawaiian residents who exhibit variation in the pronunciation of place names that have a Hawaiian origin, focusing on these linguistic variables: the glottal stop and the realization of the vowel /o/. The researchers verify whether the phonetic realizations of the Hawaiian place names are linked to the speaker’s ethnicity and/or to Hawaiian speech. After rigorous empirical research, the results of the study show evidence that both factors – the speaker’s ethnic background and his ability to speak Hawaiian – are related to the phonetic variants used. The authors argue that these phonetic achievements are indexed to social meanings of Hawaiian claims. The research also emphasizes that speakers can, through speech production, recognize the indigenous origins of these place names, as well as implicitly associate the places with the Hawaiian people and culture.

The second chapter, entitled *“Geoprosody: quantitative approaches of prosodic variation across dialects”*, authored by the French researcher Albert Rilliard, proposes, through the approach named Geoprosody, a review of the methods used to measure objective divergences between prosodic performances, typically concerning to dialectological variation. The focal point of observation is that changes in the prosodic structures of dialects are rarely addressed in these dialectometric approaches. Thus, the author discusses, focusing on Romance languages, the appropriate methods for measuring prosodic changes between

linguistic varieties, as well as the suitability of different measurements for various types of analysis, their robustness to measurement noise, their applicability and the different types of datasets.

These two chapters meet, to a certain extent, concerns that were expressed by members of the Sociolinguistics WG at the 2018 meeting and that are related to issues within the themes of subjective evaluation/ perception and dialectology/ geolinguistics.

The chapter “*Variationist Sociolinguistics and Construction Grammar: the challenges and the prospects of compatibilization*”, by Marcia dos Santos Machado Vieira and Marcos Luiz Wiedemer, departs from the assumption that it is necessary to focus on the treatment of variation in the Usage-Based Construction Grammar framework. This intent is substantiated by the articulation of this framework with that of the Variationist Sociolinguistics due to its experience (as suggested by BYBEE, in an interview made by TORRENT, 2012<sup>2</sup>, and by HOFFMANN; TROUSDALE, 2011<sup>3</sup>). In order to further advocate for this articulation, the authors focus on an effort to refine the concept of “constructional variation” and how data-based (centered on uses or records of use subjective evaluation/ perception) constructionist approaches can deal with this phenomenon. The specific focus of the chapter is on the problematization of two of the possibilities of constructional variation – the variation by independent constructions analogy/alignment and the variation of (co)lexemes compatibilization/matching to a construction slot – a novelty, in theoretical and methodological terms, in Brazil, which is beginning to draw the attention of Brazilian functionalist constructionists. Furthermore, to illustrate the configuration of these two possibilities of variation, the authors resort to constructions of verbal predication in Portuguese.

“*Perception of coda (s) variation in the speech community of Rio de Janeiro: accessing the social meaning of the back-fricative variant*”, a chapter authored by Marcelo Alexandre Silva Lopes de Melo and Christina Abreu Gomes, presents the results of a perception test which have evaluated the social meaning of the variants of the coda (s) variable and that was applied to three groups of participants belonging to two distinct social groups in the speech community of Rio de Janeiro. Based on the premise that the back-fricative

---

<sup>2</sup> TORRENT, Tiago Torrent. Entrevista. Usage-based models in linguistics: an interview with Joan Bybee. *Revista LinguiStica*, vol. 8, 1, p. 1-6, 2012.

<sup>3</sup> HOFFMANN, Thomas; TROUSDALE, Graeme. (Eds.). Special issue of Cognitive Linguistics: Variation, change and constructions in English. *Cognitive Linguistics*, vol. 22, no. 1, p. 1-23, 2011.

variant (velar/glotal) is stigmatized in several varieties of Brazilian Portuguese, the results show that not only different evaluations of coda (s) variants exist among the different groups of participants, but also that these differences are gradual. The study also demonstrates that evaluation patterns are closely related to the degree of access of the individuals from different groups to institutions that form prestige standards.

Livia Oushiro, in the chapter “*Questions and methods: pretonic midvowels in the speech of Northeastern migrants in dialectal contact*”, focus on the analysis of the patterns of variation in the height of the pretonic midvowels / e / and / o / in the speech of migrants from the Brazilian states of Paraíba and Alagoas which are residents in the city of São Paulo and discusses methods for the treatment of sociophonetic data. Results show a correlation between midvowel height and migration time. Therefore, considering the analysis based on the methods ordinarily employed, the author evaluates that the field of variationist studies in Brazil needs a change of attitude regarding the treatment of sociolinguistic data.

The text “The deletion of postonic medial vowel *in two African varieties of Portuguese: a preliminary study*”, by Danielle Kely Gomes, presents results of some analysis of the process of deletion of the medial postonic vowel, which culminates in the regularization of proparoxytones to the paroxytone pattern. In an attempt to demonstrate new evidence for the phenomenon under consideration, the author presents a comparison between data from Portuguese of São Tomé and from Portuguese of Mozambique, based on the hypothesis that proparoxytones, unnatural even for Portuguese speakers of Portuguese as L1, would be consistently regularized as paroxytones in both Portuguese varieties, as an effect of Portuguese contact with the other languages that coexist with them in both communities.

Silvia Figueiredo Brandão and Alessandra de Paula, in the chapter “*Still on the Rhotics in Portuguese in Mozambique*”, investigate which structural and social factors influence the use of R variants in the pre-vocalic position, in the speech of individuals who use Portuguese as L1 and L2, from the urban variety of Portuguese from Mozambique. The authors argue that, in the grammar of Portuguese of Mozambique, there is a single rhotic phoneme, although the results of the analysis show that tap, the dominant variant of this language, occurs in the speech of all individuals, but is significantly more frequent among those that have it as L2. Among PM L1 speakers, especially those with a higher education level, the vibrant alveolar prevails, which approximates their norm to the European norm.

The chapter written by Silvia Rodrigues Vieira, entitled “*Towards a Flexible Standard in the School Context: Contributions from Linguistic Studies*”, departs from the premise that norms practiced in various varieties/modalities need to allow for natural flexibility and adaptability by socio-communicative context. The author seeks to deepen the discussion of the establishment of the standard norm. Therefore, she presents guidelines for the treatment of pedagogical works that consider the variation. In addition, the relevance of the relationship between normative orientation and sociolinguistically grounded knowledge is highlighted.

Juliana Bertucci Barbosa and Talita de Cássia Marine, in “*Linguistic Attitudes of Master Students of PROFLETRAS units in the Triângulo Mineiro Region*”, aim to favor a discussion supported by results of a linguistic attitude test applied to teachers of Basic Education in public schools, newcomers in the Professional Master’s Degree Program in Letters (PROFLETRAS), of two units of the course in the Triângulo Mineiro region. For the test, the authors consider two variable phenomena of Brazilian Portuguese: verbal agreement (more stigmatized) and the alternation between the verbs “*ter*” and “*haver*” with the sense of existence<sup>4</sup> (less stigmatized). The purpose of the chapter is to understand sociolinguistically the linguistic attitudes of these teachers and to subsidize results for the planning of educational actions.

Each of the chapters written by Brazilian sociolinguists is followed by an extended abstract written in English. This care in the organization of this work aims to contribute to the circulation of the research practices and knowledge about the Portuguese language as wide as possible.

The organizers also offer to the readers the text “*Interview with The Sociolinguist Dinah Maria Isensee Callou*” (Professor Emeritus of UFRJ), departing from ten questions. The researcher addresses the importance of the area of Sociolinguistics in the context of linguistic research, the contributions and repercussions of this area in the national and international scenarios, its interdisciplinary character, the value of group work, as well as other subjects.

Aware of the importance of our constant interlocution/dialogue with those interested in sociolinguistic investigations from within and outside the areas of Letters and Linguistics, as well as from within and outside Brazil, the organizers of this book seek to broaden the relationship between *language research* and *society*, through the choices made in favor of the configuration of this work and

---

<sup>4</sup> Both verbal forms correspond to *there be*.

through initiatives such as the constitution of a forum<sup>5</sup> whose purpose is to make possible interlocutions that develop in the most diverse formats and that, as far as possible, meet the demands of interested parties.

We are very grateful to the authors, who were willing to contribute to the making of this book, and to the Evaluation Scientific Committee (composed of Brazilian and foreign research professors, who so readily responded to our invitation), for their productive dialogue and contributions to the important “(re) making”/ improvement involved in this process, in our work.

We especially thank the Coordination for the Improvement of the Personnel of Higher Education (CAPES) for their support. This support welcomes a proposal for the visibility of the Brazilian linguistic science that has been produced and which has contributed so much to the knowledge in the areas of Letters and Linguistics, as well as for the formation of an academic and professional staff (not only inside the country) which makes a difference in the daily lives of the communities in which this knowledge assumes features of practice(s). We also thank the UFRJ Postgraduate Program in Vernacular Letters and the UERJ Postgraduate Program in Letters and Linguistics for also embracing this idea, as well as supporting this book and the interinstitutional partnership materialized in the process of this book.

Finally, we must invite any reader interested in sociolinguistic issues to read and share the dimensions and experiences brought together in the chapters of this work, as well as to interact with the ANPOLL Sociolinguistics WG (<http://anpoll.org.br/>) and with members of the International Forum on Sociolinguistics ([www.fis2019.com.br](http://www.fis2019.com.br)).

Good reading!

*The organizers*  
*Rio de Janeiro (RJ), august 2019.*

---

<sup>5</sup> In 2019, FIS (International Forum on Sociolinguistics, [www.fis2019.com.br](http://www.fis2019.com.br)) offers the community a face-to-face event (at UFRJ's Faculty of Letters), designed by the management of the ANPOLL Sociolinguistics WG (2018-2020), which is supported by CAPES (Coordination for the Improvement of the Personnel of Higher Education), Project n. 8887.359599/2019.00, aid 865/2019, and the UFRJ Postgraduate Program in Vernacular Letters, and UERJ Postgraduate Program in Letters and Linguistics, as well as in partnership with the Federal University of Rio de Janeiro, State University of Rio de Janeiro, Università Ca' Foscari Venezia and Université Paul Valéry/Montpellier 3.



## SPEAKER ETHNICITY, LANGUAGE BACKGROUND, AND THE PRONUNCIATION OF HAWAIIAN PLACE NAMES<sup>1</sup>

*Katie Drager (University of Hawai‘i at Mānoa)<sup>2</sup>*  
*M. Joelle Kirtley (University of Hawai‘i at Mānoa)*  
*Clinton Kakela Awai (The Ohio State University)*  
*Catherine Lee (University of Hawai‘i at Mānoa)*  
*Jonny Kim (Hanyang University)*

### ABSTRACT

Residents of Hawai‘i exhibit a great deal of variation in their pronunciation of place names that have a Hawaiian origin. Using wordlist data, we investigate whether the phonetic realization of Hawaiian place names is linked to speaker ethnicity (i.e., whether the speaker has Native Hawaiian ancestry) and/or language background (i.e., whether the speaker speaks Hawaiian). We focus on two linguistic variables: the glottal stop, which is phonemic in Hawaiian, and the realization of the vowel /o/. The results provide evidence that both factors are linked with which phonetic variants are used; speakers who are Native Hawaiian and speakers who can speak at least some Hawaiian produce more Hawaiian-like realizations of the place names compared with other speakers in the study. We

---

<sup>1</sup> Acknowledgements: We would like to thank Bethany Kaleialohapau‘ole Chun Comstock for comments on earlier versions of this paper. We would also like to thank the In-Group (a sociolinguistic discussion group) for their assistance with the initial design. Of course, all errors remain our own.

<sup>2</sup> 569 Moore Hall  
1890 East-West Road  
Honolulu, HI 96822

argue that the various phonetic realizations are indexed to social meanings, and that anglicized variants can make a claim that Hawai‘i is Western and a part of the United States, whereas Hawaiian variants can reject this claim in what Herman (1999) refers to as reconquest. Further, we argue that variants can do the social work of anti-conquest and reconquest even when it is not the intent of the speaker.

## INTRODUCTION

Speakers manipulate linguistic forms to construct their social world, a process that is mitigated through sets of beliefs about language referred to as language ideologies (Woolard, 1998). Language ideologies, which include concepts such as standardness, are community-based beliefs that are constructed by the speakers themselves. The linguistic variants that tend to be associated with standardness in a society are those that tend to be used by those who have power in that society. The relationship between language and power is especially apparent in postcolonial societies because colonization results in both language contact and social change. Colonialists gained control not only of the place and the people who lived in that place, but of the people’s languages and language varieties (Woolard, 1998: 25).

In Hawai‘i, evidence of the link between power and language can be found, amongst other things, in place names. Some places have been named or renamed with words with a Western origin to appeal to tourists. For example, a place traditionally referred to as *Ohe‘o* and that was considered *kapu* (taboo) (Pukui, Elbert & Mookini, 1974: 168) began to be referred to as *Seven Sacred Pools*, reportedly a moniker made up by a nearby hotel in an effort to attract tourists<sup>3</sup>. In addition, many places have been named or renamed to honor *haole* (white) men who were directly involved in Hawai‘i’s colonization. For example, McKinley High School is named after the 25<sup>th</sup> President of the United States under whom Hawai‘i was annexed to the United States under protest. Likewise, Cook Point, a point in the northern end of Kealahou Bay, was named after the British explorer Captain Cook, whose crew members were known for committing atrocities throughout Polynesia.

In addition to renaming places, Hawaiian place names are commonly produced with anglicized realizations, where the names sound more like English and less like Hawaiian (Pukui, Elbert, & Mookini, 1974: 239; Romaine, 2002),

---

<sup>3</sup> <http://mauiuidebook.com/road-to-hana-maui/road-to-hana-sites-to-see-maui/seven-sacred-pools/>, retrieved June 13, 2019.

by Locals and non-Locals alike<sup>4</sup>. For example, the city of Honolulu is sometimes pronounced [hanəlulu] instead of the Hawaiian [honolulu]. Because of the link between colonization and the production and pronunciation of place names, many Hawaiians and non-Hawaiians feel a visceral reaction to the way a place name is pronounced, with some feeling that the anglicized variants butcher the language (Oliveira, 2009: 110-112) while others feel Hawaiian realizations sound hypercorrect and affected (Romaine, 2002: 205). We hypothesize that, as a result, the pronunciation of Hawaiian place names can be used to make claims about who has or should have power in Hawai‘i.

In this chapter, we examine the phonetic realizations of two variables, /o/ and the glottal stop, in place names that have a Hawaiian origin. We present wordlist data from Local speakers, investigating linguistic and social factors that appear to influence their phonetic realizations. The results are considered within the context of anti-conquest and reconquest (Herman, 1999). Through anti-conquest and reconquest, Local speakers can either acknowledge the place names’ indigenous origins and implicitly associate the places with the indigenous culture by pronouncing the place names as Hawaiian (i.e., reconquest) or they can make alternative claims (e.g., Hawai‘i as property of the United States and English as the rightful language of the islands) by using Anglicized pronunciations of the place names.

### *Language in Hawai‘i*

Today Hawai‘i is ethnically and linguistically diverse, a result of over two centuries of immigration, colonization, and occupation<sup>5</sup>. Prior to this extensive language contact, people in Hawai‘i spoke only Hawaiian, a Polynesian language with a (C)V syllable structure, phonemically-contrastive binary vowel length, and eight consonant phonemes, including the glottal stop (Schütz, 1981; Parker Jones 2017).

In the late 1700s and throughout the 1800s, people from all over the world began to come to Hawai‘i as visitors or immigrants, many arriving to take advantage of the fur, sandalwood, and whaling trades, and others arriving as

---

<sup>4</sup> Here, we use the term ‘Local’ to refer to people who are from Hawai‘i. It should be noted that ‘Local’ is not the same as ‘Hawaiian’, which in Hawai‘i refers to people who have Native Hawaiian ancestry.

<sup>5</sup> Settlement and contact prior to written records remains a topic of debate beyond the scope of this paper. For reference, radiocarbon data suggests a date between 1219 and 1266 AD as a date of prehistoric settlement (Wilmshurst, Hunt, Lipo; Anderson, 2011).

missionaries. Beginning in 1835, the first sugarcane plantations drew laborers primarily from China, Portugal, and several Pacific Island nations. In the late 1800s, Japanese laborers arrived in large numbers, followed by Filipino laborers in the early 1900s. Because so many different languages came into contact on the plantations, pidgins arose whereby workers could communicate with one another. In the mid-1800s, a pidgin developed that was lexified primarily by Hawaiian (Sakoda & Siegel, 2003:5). As political and economic power in Hawai'i shifted away from its original inhabitants, however, plantation workers began to use pidgins that were heavily lexified by English. Somewhere between 1900 and 1930, the primarily English lexified Hawai'i Creole, known locally as Pidgin, became the most widely used creole language (Roberts, 2004: 331; Tamura, 1993: 51), and it is still spoken in the islands today.

The Hawaiian language suffered during the late 1800s. The visitors and immigrants had brought previously unknown diseases to the islands, contributing to a drastic decline in the population of Hawaiians. In 1887, King Kalākaua was held at gunpoint and forced to sign what came to be known as the Bayonet Constitution, effectively relinquishing his power as monarch. After Kalākaua's sister, Queen Lili'uokalani, rejected the constitution upon becoming queen, a group of American businessmen – with the support of U.S. soldiers – overthrew the Hawaiian monarchy. In 1896, the Hawaiian language was banned in schools and public places<sup>6</sup> and English became the official language of Hawai'i. By the 1970s, the number of Hawaiian speakers was estimated to be around 2,000 (Warner, 2001: 135-6), though some estimates (e.g., McCarty & Lee, 2015: 346) are even lower. With the number of Hawaiian speakers dwindling, the language was in danger.

The status of Hawaiian began to change in the 1970s with the resurgence of interest in Hawaiian language and culture (Kanahale, 1979; Oliveira, 2014). Evidence of this Hawaiian Renaissance is found in several key events linked with the celebration and promotion of traditional Hawaiian life and worldviews: the revamping of the Merrie Monarch Festival in 1971, the first voyage to Tahiti of the canoe, *Hōkūle'a*, in 1976, and the establishment of the Office of Hawaiian Affairs in 1978<sup>7</sup>. The renewed interest in Hawaiian culture also extended to

---

<sup>6</sup> Act 57, sec 30 of 1896 Laws of the Republic of Hawai'i: An Act to Create an Executive Department to be known as the Department of Public Instruction; to Define its Duties and Powers: and to Repeal the Following Laws. Laws of the Republic of Hawaii passed by the Legislature at its Session, 1896. 57§30 (1896).

<sup>7</sup> OHA (Office of Hawaiian Affairs) website, accessed on 24 April 2015, from <http://www.oha.org>

language; Hawaiian language classes at the University of Hawai'i and in high schools became filled beyond capacity, and, in 1978, English and Hawaiian became co-official languages of Hawai'i. The movement continues today in what Ka'iama describes as a series of episodes where progress is "slow, steady, purposeful, and focused" (Ka'iama 2014: 112).

The Hawaiian language revitalization movement has had many successes. Indeed, for many endangered language communities – especially for those in the United States – Hawaiian "serves as a model and a symbol of hope" (Hinton, 2001:131). The number of fluent speakers, however, remains small. While Hawaiian is an official language of the state and there has been a resurgence of interest to learn Hawaiian, most people in Hawai'i – including many Hawaiians – cannot speak the language. For many people, the language lives in Hawaiian music, in select words or phrases, and in place names (Romaine, 2002: 194).

### *Place Names in Hawai'i*

Western conceptions of the relation between people and place are different than Hawaiian ones. Prior to Western influence, land was not owned. Instead, people had stewardship over the land, whereby people took care of the land, and the land took care of the people (Holmes, 2000: 44-46). Each Hawaiian island was divided into *moku*, which were divided into smaller sections called *ahupua'a*. Most *ahupua'a* extended from the mountains to the sea, and were sufficient to sustain the people living in the area. Names were given to many different aspects of the geography, such as rock formations and gulches, many with spiritual significance. Place names are prominent in genealogies and play a central role in narratives and personal introductions. They are often involved in word play and double entendre, and are commonly found in *oli* (chants), *mele* (songs), and *'olelo no'eau* (proverbs, poetical sayings, and riddles) (Pukui, 1983). Traditionally, places were rarely named after a person, and when they were, it was to denote a connection of the person to the land rather than to serve an honorific function, as is common in the West (Herman, 1999: 84).

During and following colonization, the land was commodified and divided into sections that disrupted the *ahupua'a* and, therefore, the Hawaiian way of life. Most Hawaiians ended up with little to no land and could no longer live in the places they had always known. Due to the increasing presence and power of Westerners in Hawai'i, even spaces that previously had Hawaiian names in whole or in part were given English ones. For example, *Lē'ahi* is the traditional name of the highest peak of what most people today refer to as Diamond Head (Pukui et

al., 1974: 130) and *Pu'uloa* is the traditional name for Pearl Harbor (Pukui et al., 1974: 200-201). Through the process of land seizure and (re) naming, Americans actively worked toward transforming a Hawaiian space into an American one.

After the annexation of the Hawaiian Islands by the United States, many place names were given Hawaiian names. However, this naming was done according to an American system of geographical knowledge and little understanding of the Hawaiian language. Herman (1999) argues that Americans reappropriated traditional Hawaiian names of rock formations, island divisions, and subdivisions to name streets and other symbols of American colonialism, and that, through doing so, they instantiated a form of anti-conquest in which the American colonial system offered a token of respect for Hawaiian culture while simultaneously denying the Hawaiian people sovereignty or political power. Anti-conquest is often unconscious; the people doing it generally believe they are genuinely showing respect. Even when unconscious, however, it is still damaging. When preexisting place names are reappropriated, the link between the original place and the meaning behind its name is weakened, and the story behind the link is no longer retrieved in the memories of people who say the name. Likewise, new Hawaiian-sounding names (e.g., *Hawai'i Kai* instead of the traditional *Maunaloa*) can symbolically represent changes in the ecology and physical features of the place as well as changes in economic and political power (Solomon, to appear). Herman offers "The Aloha State" as the ultimate example of anti-conquest; it includes the word *aloha* to invoke the image of "a native culture that is gracious, warm, charming, welcoming" while simultaneously and explicitly making a claim to Hawai'i as a state, property of the United States (Herman, 1999:93).

Anglicized realizations of Hawaiian place names, which are common among Hawai'i Locals, can also be viewed as examples of anti-conquest (Oliveira 2009). In some cases, speakers who produce anglicized variants may be attempting to produce them as Hawaiian, but are unable to do so because they do not speak Hawaiian. However, many speakers who use anglicized realizations appear to make no attempt to pronounce the place names as Hawaiian words; they either have no desire to learn the Hawaiian pronunciation or else they know what the Hawaiian pronunciation is, but they do not use it consistently or at all. Indeed, some Local people have a negative reaction to the Hawaiian pronunciation, interpreting Hawaiian realizations as sounding hypercorrect and affected, an interpretation which arises because the anglicized variants have become the norm (Romaine, 2002: 205). While unintentionally for most people, the anglici-

zed realizations are effectively serving the function of anti-conquest, “allowing” the Hawaiians to have Hawaiian place names while simultaneously stripping the places of one more part of their Hawaiian-ness.

Herman also discusses the concept of reconquest, in which speakers use place names to make claims about Hawai‘i as Hawaiian. The example he provides is Pukui and Elbert’s (1966) dictionary-style book *Place Names of Hawai‘i*, which lists Hawaiian place names and provides explanations and translations. This book – and the expanded edition that followed (Pukui et al., 1974) – serve as excellent examples of how people can reclaim their connections to the land by returning to a precolonial understanding of places and the important role those places play in Hawaiian life, culture, and worldview. Other examples of reclaiming the connections to the land include the installation of signs denoting pre-colonial geographic divisions (Chang, 2014) and the ongoing movement to reclaim sacred places such as *Kaho‘olawe* and *Mauna Kea*<sup>8</sup>.

As with anti-conquest, pronunciation can contribute to efforts of reconquest. Hawaiian pronunciations of Hawaiian place names can allow a speaker to make a claim about Hawai‘i as Hawaiian, a phrase that likely has different meanings for different people. First, some speakers might use the Hawaiian pronunciations to demonstrate that they treat and view the land in ways that are consistent with traditional Hawaiian values. Secondly, Hawaiian pronunciations could be used to recognize the historical traditions of the place and show respect to the Hawaiian people. In addition, Hawaiian pronunciations might be used to express a desire for Hawai‘i to gain independence from the United States, where the pronunciations of place names are a form of resistance (cf. Kearns & Berg 2002: 287). Hawaiian pronunciations of place names might index all of these meanings or only select ones, depending on the speaker and the situation. In this paper, we use the umbrella term *Hawai‘i as Hawaiian* to refer to the multitude of social meanings that might be indexed. We argue that by using more Hawaiian-like pronunciations and pushing toward a better understanding of the meanings behind place names, Hawaiians are engaging in reconquest of their homeland. Thus, we argue that anglicized and Hawaiian realizations of Hawaiian place names are examples of anti-conquest and reconquest, respectively: speakers can make claims about

---

<sup>8</sup> There is a long-standing and on-going dispute between astronomers who use the mountain, Mauna Kea, as a viewing space and Native Hawaiian activists in the Kū Kia‘i Mauna movement, for whom the mountain is one of the most sacred places in Hawai‘i. Following the arrests of protesters in the spring of 2015, the mountain and debate received worldwide attention (Herman, 2015).

their stance on Hawai'i as Hawaiian depending on the phonetic variants they use to pronounce Hawaiian place names.

While there are many ways in which speakers anglicize Hawaiian words, we focus on two in this paper: alternative realizations of the vowel /o/ and the omission of the glottal stop. The glottal stop is a consonant in Hawaiian, so pairs such as *kou* ([kou] 'your') and *ko'u* ([koʔu] 'my') have different meanings and different numbers of syllables (Schütz, 1994: 143) although there can be variation in some lexical items, such as *kāua* 'you and me/I', which can be pronounced [ka:ua] or [ka'ua] (NeSmith 2005). Phonetically, the sound is variably realized as creak or a stop (Parker Jones 2017), with evidence that it is usually realized as creak in spontaneous speech (Drager, Chun Comstock & Kneubuhl, 2017: 77). Early missionaries who developed the writing system for Hawaiian did not include the glottal stop in the orthography. Since then, scholars have recognized the glottal stop as a consonant, and some members of the Hawaiian speaking community have pushed for it to be represented as an inverted apostrophe, referred to as an 'okina<sup>9</sup>, while others – particularly those who learned Hawaiian through an unbroken chain of intergenerational transmission – prefer the words to be written without an 'okina. A great deal of debate has centered around the glottal stop (Romaine, 2002) and, consequently, people throughout the islands – regardless of their ethnicity or language background – are aware that it exists and that it can be represented with an 'okina.

A sound that has received much less attention is /o/, as in the word *Honolulu*, which native speakers of Hawaiian realize as a rounded close-mid back vowel [o] that is monophthongal. In some Locals' anglicized productions of Hawaiian words, the vowel [o] is reduced to a schwa, realized as a diphthong, or realized as an entirely different vowel (e.g., /a/ in [hanəlulu]). The non-Hawaiian realizations of this vowel are especially intriguing because, like in Hawaiian, this vowel is realized as back and monophthongal in both Hawai'i English (Kirtley, Grama, Drager & Simpson, 2016) and Pidgin (Grama, 2015), the two most widely spoken language varieties in Hawai'i. This suggests that the realizations are not anglicized according to the closest vowel in the speakers' variety of English or the English-lexified creole but that the speakers are either approximating diphthongal variants found in the continental United States or are changing the vowel identity altogether. While many instructors of beginner Hawaiian language classes explicitly instruct their students to produce [o] in place names like Honolulu, this variable has received much less attention in the media, is not as marked in the

---

<sup>9</sup> Older names for the sound and/or symbol include *kai'i*, *kai'i'i*, and 'u'ina (Schütz, 1994: 146).

spelling, and is overall less noticed than the glottal stop.<sup>10</sup> As such, speakers are likely to pay less attention to /o/ than the glottal stop in tasks where the glottal stop is written and, therefore, speakers who use anglicized realizations in spontaneous speech may be less likely to shift to Hawaiian realizations during a wordlist task such as that used in the current study.

We hypothesized that a speaker's ability to speak Hawaiian would be linked with whether or not they produced the more Hawaiian-like realizations of the glottal stop and /o/. Because of the potential for the pronunciation of place names to play a role in anti-conquest and reconquest, we also hypothesized that speakers who are Native Hawaiian would produce more tokens with Hawaiian realizations than the non-Hawaiian speakers.

## METHODS

### *Data Collection*

The data analyzed for this paper are taken from SOLIS, a multi-language archive of conversation and interview data that is housed at the University of Hawai'i at Manoa and is overseen by the first author. Speakers are recruited from the community by word of mouth and from the university campus through a participant pool, and several of the authors of this paper are among the interviewers. As part of the interviews conducted for inclusion in SOLIS, participants read several wordlists, one of which contains names of places from around Hawai'i. These names have a Hawaiian origin and demonstrate variation in how they are realized. The words were read in isolation and in a fixed order. The results reported in this paper focus entirely on the place names from this wordlist data and, specifically, on the analysis of two linguistic variables: /o/ and the glottal stop. The place names subject to the analysis are given in Table 1.

When designing the wordlist, we chose to represent the *'okina* (glottal stop) and *kahakō* (macron indicating vowel length) in the wordlist because wordlists elicit careful, self-conscious speech, and we wanted to examine which speakers would choose to produce the glottal stop when given every opportunity to do so. Speakers were instructed to read the place names as they would normally say them, but we anticipated that at least some speakers would shift toward a

---

<sup>10</sup> It is our belief that this variable is not entirely below the level of consciousness when in place names; if someone were asked to imitate a White American, they would probably produce alternative variants of /o/. However, /o/ is not as marked as the glottal stop.

more Hawaiian-like pronunciation than they use in everyday speech. This is not an undesired effect because we were explicitly interested in the ways in which speakers might use these variables to do social work, such as constructing their ethnic identities or taking a stance on Hawai‘i as Hawaiian. Future work will examine variation of the variables in spontaneous speech.

**Table 1:** Hawaiian place names collected and analyzed, for the glottal stop, /o/, or both variables.

place name	analyzed for
Hawai‘i	glottal stop
Hawai‘i Kai	glottal stop
Kaua‘i	glottal stop
Lā‘ie	glottal stop
Lāna‘i	glottal stop
Līhu‘e	glottal stop
Ni‘ihau	glottal stop
Wai‘alae	glottal stop
Wai‘anae	glottal stop
Kaho‘olawe	both
Kāne‘ohe	both
O‘ahu	both
Hilo	o
Honolulu	o
Kona	o
Moloka‘i	o

### *The Speakers*

The analysis was conducted on wordlist data from 55 participants, all of whom were born and raised in Hawai‘i and took part in the interviews between 2009-2014. Demographic information about the participants (e.g., whether they speak Hawaiian) was gleaned from a combination of information provided on an information sheet and during the interview. On the information sheet, participants were asked to list their ethnicities as well as indicate information about their age and gender.

People who reported speaking any Hawaiian were categorized as Hawaiian speakers for the purposes of this study. In addition, participants were categorized depending on whether they listed Hawaiian among their ethnicities.

Table 2 shows the number of participants by their ethnicity and whether they could speak Hawaiian. The table also divides speakers into categories based on their self-reported age and gender: younger men and women (YM and YW) were born between 1975 and 1995, which means they were born during or after the Hawaiian Renaissance of the 1970s. In contrast, the older participants (OM and OW) were born between 1924 and 1969, prior to when the Hawaiian Renaissance had really blossomed. The mean and median ages of each group are shown in Table 3.

**Table 2:** Summary of participant demographics by their ability to speak Hawaiian and their ethnicity, shown separately for by the participant’s age-gender group: younger women (YW), older women (OW), younger men (YM), and older men (OM).

Language background	Hawaiian ethnicity	YW	YM	OW	OM	Total
speaks Hawaiian	Hawaiian	6	6	1	2	15
	non-Hawaiian	2	0	0	0	2
cannot speak Hawaiian	Hawaiian	8	3	6	1	18
	non-Hawaiian	5	6	5	4	20
Total		21	15	12	7	55

**Table 3:** The minimum, maximum, mean, and median ages of participants at the time of participation, by gender and age group. Median ages are shown in parentheses.

age/gender group	min	mean (median)	max	N
YW	18	21 (22)	24	21
YM	18	22 (21)	35	15
OW	40	66 (64)	91	12
OM	48	62 (58)	76	7

### *Auditory Analysis*

Two listeners conducted auditory analysis on each variable, categorizing each token as one of five realizations. All tokens for which there was disagree-

ment were then re-coded by the first author using auditory analysis. Two tokens were removed prior to analysis because there was disagreement among all three coders. Tokens with background noise that obstructed the target sound were also not included in the analysis.

The glottal stop was initially coded as being realized as: a stop, creak, mild creak, no glottal quality, and other (e.g., a small number of participants produced Ni‘ihau as [nihiʔau]). For simplicity, these more detailed categories were condensed into two: any glottal quality versus none or other. Thus, realizations coded as having either a glottal stop, creak, or mild creak are treated as realizations of the glottal stop. This division makes the most sense given that many speakers of Hawaiian realize the phoneme as creak in at least some phonological environments (Parker Jones 2017: 104-5). Eleven tokens were removed from the 660 collected tokens (55 speakers x 12 words), and a total of 649 tokens were used for the analysis of the glottal stop presented herein.

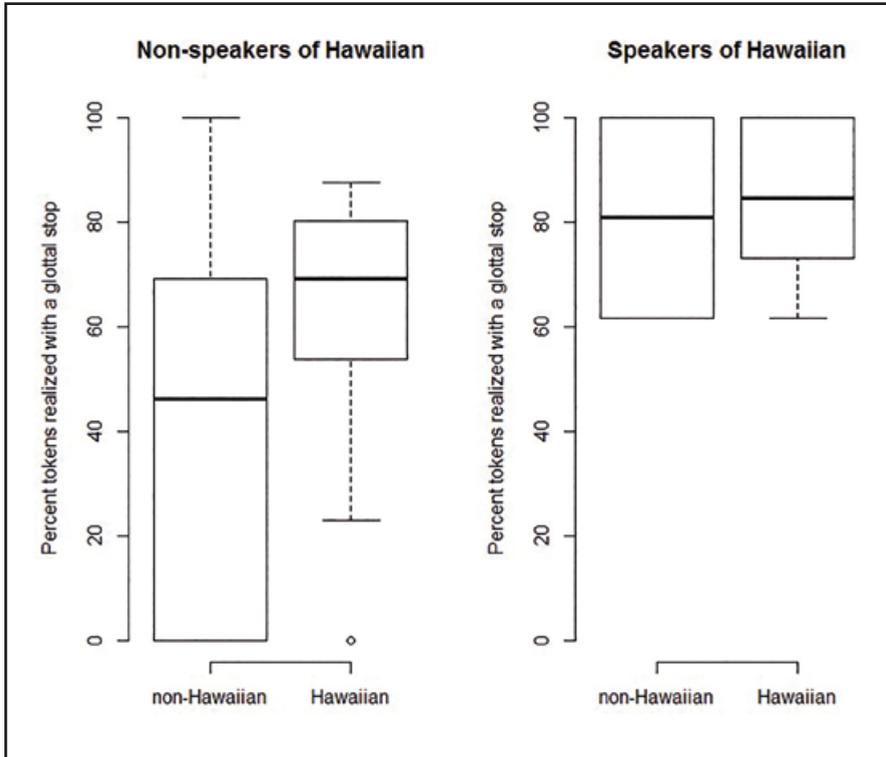
/o/ was initially coded into one of five categories: monophthongal, slightly diphthongal, very diphthongal, schwa, and any other vowel. For the analysis presented here, these were collapsed into a binary distinction between whether a token was realized as the monophthongal back vowel [o] versus any other realization. Words containing more than one instance of /o/ were coded as “other” if either /o/ was realized as anything other than [o]. 14 tokens were removed, and a total of 371 tokens were analyzed for /o/.

## RESULTS

### *Glottal Stop*

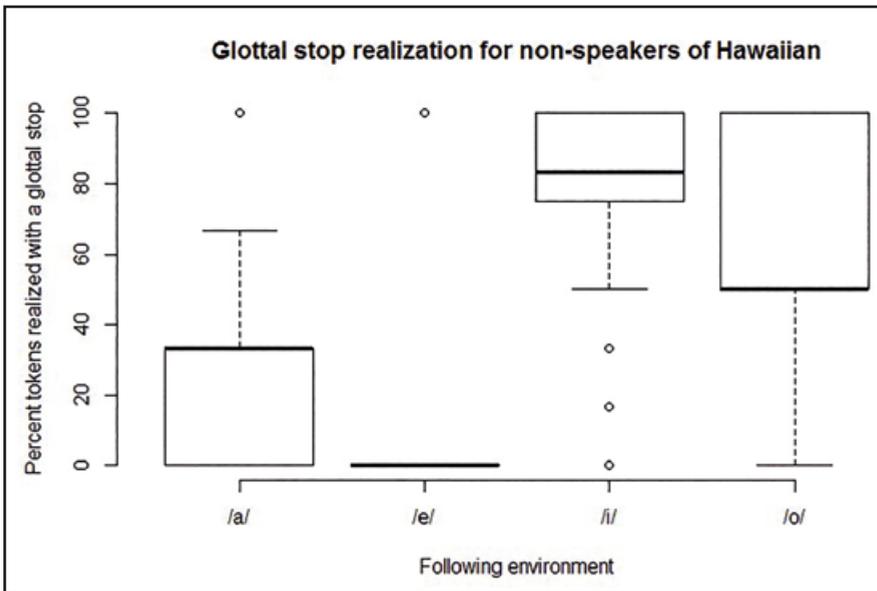
For the production of the glottal stop, we observed a difference between the participants who speak some Hawaiian and those who do not; Hawaiian speakers realized the glottal stop in 83% of tokens, whereas participants who do not speak Hawaiian produced it in only 58% of tokens. This difference is evident in Figure 1. Also evident in Figure 1 is that, among those who do not speak Hawaiian, ethnicity is linked with rates of glottal stop production; speakers who are Native Hawaiian produced a higher percentage of tokens with the glottal stop.

**Figure 1:** Box and whisker plots showing the percentage of each speaker's tokens in which the glottal stop was realized. The speakers' percentages are shown separately for those who speak some Hawaiian (right panel) and those who do not (left panel), and for whether the speaker is Native Hawaiian or not.



The intraspeaker variation observed appears to be partially conditioned by the lexical item, and may be linked with the following phonological environment. As shown in Figure 2, the glottal stop is most likely to be realized in the data when preceding the high front vowel /i/ and less likely to be realized when preceding /a/. Since only a small number of lexical items are included in each phonological environment in this study, further work is required to determine whether this tendency is generalizable to place names not included in our wordlist.

**Figure 2:** Box and whisker plot showing the percentage of each speaker's tokens followed by /a/, /e/, /i/, and /o/ in which the glottal stop was realized.



To test the statistical significance of the trends evident in Figures 1 and 2, binary logistic regression models were fit to the data using the lme4 package (Bates, Maechler, Bolker & Walker, 2015) in R (R Core Team, 2014). The binary variable of whether or not the glottal stop was realized (Glottal.presence, henceforth) was treated as the dependent variable, and both speaker and item were included as random intercepts. Random slopes were not included due to non-convergence as a result of data sparsity. Tested in the model were the speaker's self-reported ability to speak at least some Hawaiian (Speak.Hawaiian), whether the speaker was Native Hawaiian (Native.Hawaiian), following environment (Following.vowel), and the speaker's age and gender. Due to sparseness of the data and unbalanced cells, the model is overfit with both social factors included, so the model was trimmed of Native.Hawaiian, as it was the weaker predictor. Neither age nor gender reached significance in the model.

**Table 4:** Output of logistic regression model fit to the binary dependent variable of whether or not the glottal stop was realized: model = glmer(Glottal.presence ~ Following.vowel + Speak.Hawaiian + (1|speaker) + (1|item), data=Glottal.data, family=binomial).

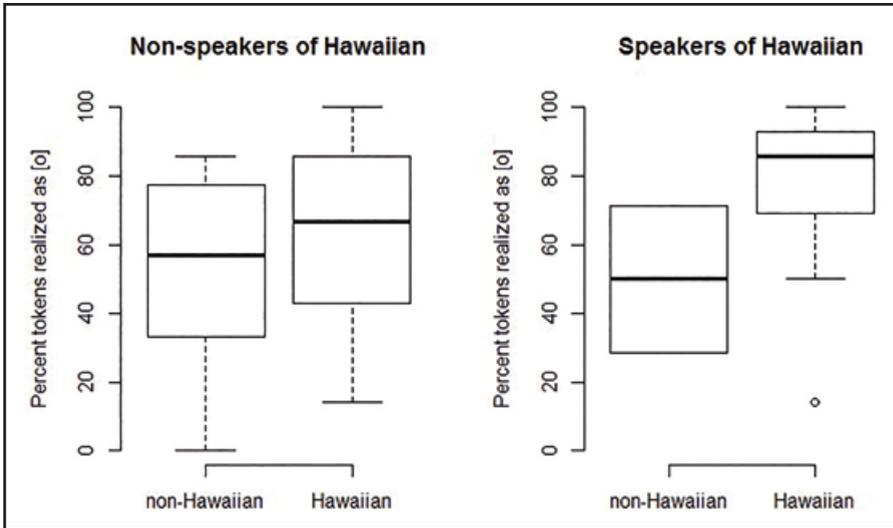
	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	-1.5549	0.6145	-2.53	0.0114
Following.vowel = e	-0.5878	1.0876	-0.541	0.5889
Following.vowel = i	3.6075	0.7056	5.113	<.0001
Following.vowel = o	2.3815	0.8849	2.691	0.0071
Speak.Hawaiian = y	2.4234	0.5819	4.165	<.0001

The output of the model is shown in Table 4. Following.vowel was found to have a significant effect, whereby a glottal stop is more likely to be realized when followed by either /o/ ( $p<.01$ ) or /i/ ( $p<.0001$ ) compared with /a/. The glottal stop is also significantly more likely to be realized if the participant can speak at least some Hawaiian ( $p<.0001$ ).

### *The realization of /o/*

The speakers produced a variety of different realizations for /o/. The most common anglicized realizations were [a] and schwa, as in [hanəlulu] for Honolulu. The distribution of [o] varied by item. Among the words tested, *Kona* and *Kaho'olawe* were most often realized with the back monophthongal variants, with even the non-Hawaiian speakers pronouncing /o/ as [o] in these words over 80% of the time. The word *Honolulu* had the largest number of alternative forms; non-speakers of Hawaiian realized /o/ as [o] only 32% of the time. For most items, the most common alternative realization was a vowel other than [o] (most often [a], [ʌ], or [ə]) but in the word *Hilo*, the vowel was often realized as diphthongal.

**Figure 3:** Box and whisker plots showing the percentage of each speaker’s tokens realized as [o]. The speakers’ percentages are shown separately for those who speak some Hawaiian (right panel) and those who do not (left panel), and for whether the speaker is Native Hawaiian or not.



The results indicate that speaker ethnicity played a role in whether /o/ was realized as [o] or not. As shown in Figure 3, speakers who are Hawaiian were more likely to realize /o/ as [o] compared to speakers who are not Hawaiian, regardless of whether or not they could speak Hawaiian. Among Hawaiians, there may also be a link between /o/ realization and whether they speak Hawaiian or not. However, a larger and more balanced sample is needed to confirm this.

Binary logistic regression models were fit to the binary variable of whether or not the /o/ was realized as [o] (O.realization). Speaker and item were included as random intercepts. Tested in the model were Speak.Hawaiian, Native.Hawaiian, as well as the speakers’ age and gender categories. Due to sparseness of the data, the model would not converge with all social factors included, so only the strongest predictor, Native.Hawaiian, was retained in the model.

**Table 5:** Output of logistic regression model fit to the binary dependent variable of whether or not the /o/ was realized as [o]: model = glmer(O.realization ~ Native.Hawaiian + (1|speaker) + (1|item), data=O.data, family=binomial)

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.0937	0.6192	0.151	0.8797
Native Hawaiian = y	1.3650	0.5110	2.671	0.0076

The output of the model, shown in Table 5, indicates that the tendency for speakers who are Hawaiian to produce /o/ as [o] more often than speakers who are not Hawaiian is statistically significant ( $p < .01$ ).

## DISCUSSION

The results provide evidence that both linguistic and social factors influence the realization of place names that have a Hawaiian origin. The realization of the glottal stop was found to be most closely linked with language background and the following phonological environment, whereas /o/ realization was found to be most closely related to the speaker's ethnicity.

Why have we observed a difference in /o/ realization across ethnicities? One possible interpretation is that Hawaiians have more exposure to Hawaiian pronunciation of the place names as a result of their family histories (e.g., grandparents who speak Hawaiian); the pronunciation of place names may be passed down even if the language as a whole is not. Those with Hawaiian ancestry may also use a greater number of Hawaiian words and phrases in day-to-day speech and have more exposure to Hawaiian through, for example, Hawaiian festivals, gatherings, and school-related events. Another possibility is that these realizations are actively used to construct the speakers' ethnicities, which may in turn be linked with the speakers' stances on the position of Hawai'i as Hawaiian. It seems most likely to us that it is some combination of these interpretations: people with Hawaiian ancestry are more likely to have had exposure to the Hawaiian pronunciations of these place names, and then they actively use these variants to construct their Hawaiian-ness in interaction and to take stances on Hawai'i as Hawaiian. For many years, Hawaiians were made to feel ashamed of their culture by those in power (Silva, 2004; Marshall, 2006) and were blamed for the decimation of their own culture (see e.g., Kroeber 1921: 130). However, the Hawaiian Renaissance, the sovereignty movement, and related movements to reclaim sacred spaces have fostered and continue to foster cultural pride amongst those with Hawaiian ancestry; Hawaiian culture is a key part of what it means to be Hawaiian (Osorio 2001: 362). Using Hawaiian pronunciations of Hawaiian place names can demonstrate pride in one's ethnic heritage and can serve to underscore a speaker's sense of place and cultural identity.

Why then have we not observed an equivalent difference across ethnicities for the glottal stop? We believe this is a combination of two factors. The first is that we observed a strong effect of Hawaiian language background, whereby people who had studied Hawaiian were more likely to produce the glottal

stop, regardless of their ethnicity, and their production of the glottal stop is near ceiling. The second is that Locals have more metalinguistic awareness about the glottal stop than about /o/; most Locals are aware that the glottal stop is a sound in Hawaiian, but there is little metalinguistic discussion around /o/, and the difference is especially notable during our task since the glottal stop was represented in the orthography but people spell the place names with the letter <o> even when they do not pronounce the sound as [o]. Taken together, this would mean that any speakers who were predisposed to produce the variables with Hawaiian realizations would likely try to produce the glottal stop, and those who had studied Hawaiian were especially capable of producing it as a result of their training. In contrast, /o/ was likely realized as [o] primarily by the speakers who tend to produce it as [o] in the place names in spontaneous speech.

The written cue for the glottal stop provided all of the speakers with the opportunity to produce it in the place names. So why did some speakers produce zero to few instances of the glottal stop? For some speakers, not realizing the glottal stop could be a reflection of their inflexibility in changing the pronunciations of place names from those that they learned when they were young. For other speakers, it may be a reflection of their resistance to the Hawaiian renaissance and the reconquest of place names. Further research is required to explore these different possibilities, but it is important to keep in mind that, while the use of the anglicized variants may sometimes be unintentional, they are doing the work of anti-conquest through maintaining the status quo as Western and reinforcing many Americans' assertion that Hawai'i is Western and a part of the United States.

## CONCLUSION

We have demonstrated that there is considerable variation in the way that Hawai'i Locals pronounce place names of Hawaiian origin and that this variation is linked with both a speaker's ethnicity and their ability to speak Hawaiian. We have further suggested that this variation could meaningfully reflect speakers' attitudes towards Hawaiians' claims to place. Through producing more Hawaiian-like realizations, speakers can acknowledge the indigenous origins of these place names as well as implicitly associate the places with Hawaiian people and culture. Future work that includes an analysis of more spontaneous contexts would shed light on how speakers use phonetic variation in place names to perform acts of identity and solidarity.

## REFERENCES

- Chang, C. (2014). Ko'olauloa welcomes new boundary sign. *Midweek*, 18 June 2014, retrieved on 20 March 2015, from <http://www.midweek.com/hawaii-community-news/windward/windward-oahu-coverstory/koolauloa-welcomes-new-boundary-sign/>
- Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. <doi:10.18637/jss.v067.i01>.
- Drager, K., Chun Comstock, B.K., & Kneubuhl, H.P. (2017). He nui nā ala e hiki aku ai: Factors influencing phonetic variation in the Hawaiian word kēia. In Hildebrandt, K.A., Jany, C. & Silva, W., editors, *Language Documentation & Conservation Special Publication No. 13*, pages 65–93.
- Grama, J. (2015). Variation and change in Hawai'i Creole vowels. Unpublished doctoral dissertation, University of Hawai'i at Mānoa.
- Herman, D. (2015). The heart of the Hawaiian people's argument against the telescope on Mauna Kea. *Smithsonian Magazine*, retrieved on 20 March, from <http://www.smithsonianmag.com/ist/?next=/smithsonian-institution/heart-hawaiian-people-arguments-arguments-against-telescope-mauna-kea-180955057/>
- Herman, R.D.K. (1999). The Aloha State: Place names and the anti-conquest of Hawai'i. *Annals of the Association of American Geographers*, 89(1), 76-102.
- Hinton, L. (2001). An introduction to the Hawaiian language. In Hinton, L. and Hale, K., editors, *The Green Book of Language Revitalization in Practice*, pages 129–132. Academic Press, San Diego.
- Holmes, L. (2000). Heart knowledge, blood memory, and the voice of the land: Implications of research among Hawaiian Elders. In Sefa, G.J, Hall, B.L., and Goldin Rosenberg, D., editors, *Indigenous knowledges in global contexts: Multiple readings of our world*, pages 37-53.
- Ka'iama, M. (2014). Kū i ka pono: The movement continues. In N. Good-year-Ka'ōpua, I. Hussey, & E. K. Wright (Eds.) *A Nation Rising: Hawaiian*

*movements for life, land, and sovereignty* (pp. 98-114). Durham & London: Duke University Press.

Kanahele, G.S. (1979). The Hawaiian Renaissance. Polynesian Voyaging Society Archives: Primary Source Documents 1978-1984, May 1979, retrieved on 20 March 2014, from <http://kapalama.ksbe.edu/archives/PVSA/primary%202/79%20kanahele/kanahele.htm> Kearns, R.A., & Berg, L.D. (2002). Proclaiming place: Towards a geography of place name pronunciation, *Social and Cultural Geography*, 3(3), 283-302.

Kirtley, M.J., Grama, J., Drager, K., & Simpson, S. (2016). An acoustic analysis of the vowels of Hawai'i English, *Journal of the International Phonetic Association* 46(1): 79-97. <https://doi.org/10.1017/S0025100315000456> Kroeber, A.L. (1921). Observations on the Anthropology of Hawaii. *American Anthropologist*, 23: 129-138.

Marshall, W.E. (2006). Remembering Hawaiian, transforming shame. *Anthropology and Humanism*, 31(2), 185-200.

McCarty, T.L., & Lee, T.S. (2015). The role of schools in Native American language and culture revitalization: A vision of linguistic and educational sovereignty. In W.J. Jacob, S.Y. Cheng, & M.K. Porter (Eds.), *Indigenous education: language, culture, and identity* (pp. 341-360). New York: Springer.

NeSmith, R. K. (2005). Tūtū's Hawaiian and the emergence of a Neo Hawaiian language, *Ōiwi: A native Hawaiian journal*, 3, 135-152.

Oliveira, K. R. K. N. (2009). Wahi a kahiko: Place names as vehicles of ancestral memory. *AlterNative: An International Journal of Indigenous Peoples*, 5 (2), 100-115).

Oliveira, K. R. K. N. (2014). E ola mau ka 'Ōlelo Hawai'i. In N. Good-year-Ka'ōpua, I. Hussey, & E. K. Wright (Eds.) *A Nation Rising: Hawaiian movements for life, land, and sovereignty* (pp. 78-85). Durham & London: Duke University Press.

Osorio, J.K. (2001) “What kine Hawaiian are you?” A mo‘olelo about nationhood, race, history, and the contemporary sovereignty movement in Hawai‘i. *The Contemporary Pacific* 13(2): 359-379.

Parker Jones, ‘Ō. (2017). Hawaiian. *Journal of the International Phonetic Association* (2018) 48/1: 103-115.

Pukui, M. K. (1983). *‘Ōlelo No ‘eau, Hawaiian proverbs and poetical sayings*. Honolulu, HI: Bishop Museum Press.

Pukui, M.K., & Elbert, S. (1966). *Place names of Hawaii and supplement to the third edition of the Hawaiian-English dictionary*. Honolulu: University of Hawai‘i Press.

Pukui, M.K., Elbert, S., & Mookini, E.T. (1974). *Place names of Hawai‘i* (2nd ed.). Honolulu: university of Hawai‘i Press.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Roberts, S. (2004). The role of style and identity in the development of Hawaiian Creole. In G. Escure, & A. Schwegler (Eds.), *Creoles, contact, and language change: linguistic and social implications* (pp. 1-20). Philadelphia, PA: John Benjamins.

Romaine, S. (2002). Signs of identity, signs of discord: glottal goofs and the green grocer’s glottal in debates on Hawaiian orthography. *Journal of Linguistic Anthropology* 12(2), 189-224.

Sakoda, K., & Siegel, J. (2003). *Pidgin Grammar: An introduction to the creole language of Hawai‘i*. Honolulu: Bess Press.

Schütz, A.J. (1981). A reanalysis of the Hawaiian vowel system, *Oceanic Linguistics* 20(1), 1-43.

Schütz, A.J. (1994). *The voices of Eden: A history of Hawaiian language studies*. Honolulu: University of Hawai'i Press.

Silva, N.K. (2004). *Aloha betrayed: Native Hawaiian resistance to American Colonialism*. Durham: Duke University Press.

Solomon, H.N. (to appear). Maunaloa: Shifting nomenclatures and spatial reconfiguration.

Tamura, E.H. (1993). The English-only effort, the anti-Japanese campaign, and language acquisition in the education of Japanese Americans in Hawai'i, 1915-40. *History of Education Quarterly* 33(1), 37-58.

Warner, S. N. (2001). The movement to revitalize Hawaiian language and culture. In L. Hinton & K. Hale (Eds.), *The green book of language revitalization in practice* (pp. 133-144). San Diego, CA: Academic Press.

Wilmshurst, J.M., T.L. Hunt, C.P. Lipo, & A.J. Anderson. (2011) High-precision radiocarbon dating shows a recent and rapid initial human colonization of East Polynesia. *Proceedings of the National Academy of Sciences*, 108(5):1815-1820.

Woolard, K.A. (1998) Introduction: Language ideology as a field of inquiry. In B.B. Schieffelin, K.A. Woolard, & P.V. Kroskrity (Eds.), *Language ideologies: Practice and theory*, (pp. 3-47). Oxford: Oxford University Press.

## **GЕOPROSODY**

### QUANTITATIVE APPROACHES OF PROSODIC VARIATION ACROSS DIALECTS

*Albert Rilliard (LIMSI, CNRS, Université Paris Saclay; UFRJ)*

#### **ABSTRACT**

This chapter aims at summarizing approaches to dialectometry proposed in the literature since Seguy and Goebel pioneering works, as well as on the basis of metrics used in sociolinguistics. Based on the literature review, it discusses these different metrics under a still under studied angle: their application to prosodic variation, especially for studying dialectological changes. Possible objective measures are regrouped in two broad categories, depending on whether they use descriptions based on melodic contours or on feature sets. Proposals for use of such metrics are made, in relation to existing phonological and phonetic approaches to prosody description. Examples of application of these metrics are then given on a small subset, and their output discussed and reconciled. It shows several approaches are available, that basically give convergent objective information, and can be applied to studies based on different theoretical backgrounds.

## INTRODUCTION

Since the pioneering work of Séguy (1971), the idea of quantitative measurements of linguistic differences across dialects has flourished. Séguy was working on lexical differences, and pleads for a computational approach due to the complexity of the task. Building on this founding stone, Goebel (e.g. 1982, 2006, 2010) laid the bases of computational dialectometry, introducing notably measures of divergences between linguistic varieties based on similarity and dissimilarity measurements (a similarity being easily converted into a dissimilarity, we'll use mostly the term "divergence"), their grouping into "distance matrices" the content of which is then represented on choropleth maps (or variations of it representing for example boundaries, see INOUE 1996). Representations of language variation may be made in relation to a reference point, so to display the changes occurring in language from a given point of view. Maps may also present main zones characterized by stable feature sets, by enhancing the transition regions between main dialectal varieties (HEERINGA; NERBONNE 2013). These dialectometric approaches target various levels of linguistic description, including its lexical, diachronic, orthographic, phonetic, phonologic, or morpho-syntactic aspects (see GOEBL, 2003; HEERINGA, GOOSKENS, 2003; PEIRSMAN et al. 2010 for different examples). Changes in the prosodic structures of dialects are seldom addressed in these dialectometric approaches. Even a reference work such as Chambers & Trudgill's "dialectology" (2004), or the review article by Wieling & Nerbonne (2015) on "advances in dialectometry", does not mention the terms "prosody" and "intonation". Prosodic changes were used in several works, but mostly within perceptual approaches (GOOSKENS 1997; MASE 1999; KUIPER 1999), a fact that underlines its importance in the reception of language variation, and the significance it may have to take it into account for dialectometry. An early dialectometric work on prosody was proposed by Gooskens & Heeringa (2006), who took into account stress and tonemes to calculate a "prosodic distance".

Changes within a language are not only linked to the geographic distance across population speaking dialectal varieties. Factors such as education, exposure to audio-visual media, self-affirmation and cultural empowerment, may have important effects on language use (CHAMBERS & TRUDGILL 2004), without being necessarily best analysed on the basis of geographic spread. Dialects may also be used diastatically by competent speakers, according to the communication situations – a view that is coherent with Séguy's claim that dialects serve both to communicate and to differentiate the speaker from other social groups.

Whatever the complexity of the various sources of variation, similar techniques designed to quantify linguistic differences may equally be applied; this shall not be the case for representation methods – but several proposals exist: see for example Inoue’s glottograms (2016) used to represent age and geographic spread of (lexical) changes in language use.

The central notions of dialectometry are: (i) measures of differences or similarities across language varieties, for one or several linguistic functions; (ii) agglomeration and clustering algorithms to organize large datasets according to their proximity/differences; (iii) tools for the representation of measured variations, across time, space and social groups. From these methods, the first ones are the most dependent on the types of data one wants to analyse, the two others being mostly similar whatever the linguistic analysis and mostly dependant on datasets and research aims.

Measuring of differences between linguistic features has often been based on several kinds of string metrics (typically Hamming or Levenshtein distances: SEGUY 1971; GOOSKENS & HEERINGA 2004). It basically consists in counting the number of differences between two series of labels representing features (letters, phonemes, phonetic features, etc.) that may change along the dialectal continuum. Other measures, notably inspired by sociolinguistics methodology, are based on frequency differences between kinds: the relative frequencies of studied items (lexical, grammatical, phonetic, etc.) may be used as a metric to evaluate differences in language usage (in these cases, the frequency of a phenomenon is used rather a divergence based on the comparison between two versions of the phenomenon). Such approach may compare different varieties for their relative use of a given phenomenon (e.g. WEINER et al. 1983), but more recent method may also apply the approach mixing the social and geographic dimensions (WIELING et al. 2014). The approach proposed by Speelman & Geeraerts (2008) groups lexical variants of the same concept under one “profile”, so to take into account a finer description of lexical use, which is not necessarily binary, as it may be seen in atlases or other approaches; profiles are then compared across language varieties or across time, based on several metrics.

The next step, once speakers’ productions have been compared for their varying linguistic characteristics, consists in agglomerating the divergences into matrices representing the comparisons between all pairs of the considered entities (speaker, village, time, etc.) in the dataset. These matrices (often referred to as “distance matrices”) are then subjected to a multidimensional analysis, typically a multidimensional scaling (MDS; see Baayen, 2008, for details): the output of

these methods will give a reduced set of abstract dimensions that best represent the variation between the individuals represented in the input matrix. In the case of MDS, the solution is forced on two dimensions, a practice that is particularly interesting to study variations that are supposed to take place along geographic space. The output of the multidimensional analysis allows extrapolating distances (i.e. linguistic distances) between the compared items (ROMNEY et al. 2000). Two points (i.e. the items of the input) that are the closest according to this measurement may be supposed to share most of their linguistic competences; two points that are far removed shall have marked distinctive linguistic practices. It is possible to regroup these points according to these distances, using whatever clustering method best fit the purpose of the study. Hierarchical clustering approaches are often used (GOEBL 2003, HEERINGA & GOOSKENS 2003), and produce dendrogram representations of the distances between items, allowing a simple comparison between groups.

Then, the divergences between items or groups may be represented so to allow meaningful analysis of the results. Representations in the case of dialectology are often based on maps. Goebel (2006) heralded the use of similarity maps: first Delaunay-Voronoi polygonal structures allow partitioning map with spread inquiry points into a continuum, each polygon can then be filled according to the distance the related point has with a reference point; other maps may enhance the boundaries between inquiry points (so to mark strong dialectal variations), or the link between two point (so to marks zones of similarity). Such cartography has been evolved with the progress of computer graphic technology (Wieling & Nerbonne 2015).

The aim of this chapter is to discuss methods suited to measure prosodic changes across linguistic varieties. The suitability of the different measures to several types of analyses, their robustness to measurement noise, their applicability to different types of datasets will be discussed. The discussion as well as the works presented will have a focus on Romance languages, especially on the application part; specific aspects linked to such kind of prosodic measurements for tonal languages, particularly, will not be addressed. Existing solutions will be presented and challenges for better approaches and representations of such measurements will be presented.

## DESCRIPTION OF PROSODIC CHANGES

As prosody may be viewed as the domain of gradual variation, as opposed to phonemic categories, it has long been approached using different means to

control its continual changes. One of the most fruitful such approach may be found in the theoretical and practical works made at IPO that culminates in t'Hart et al. book (1990). The process of stylisation that is advocated by these researchers, simplifying the fundamental frequency ( $F_0$ ) curve as a series of straight lines so to remove microprosodic and other involuntary changes from pitch contours, had a main influence on prosodic description works. The first step is linked with the calculus of a close-copy stylization that carries phonetic variations; the second one consists in creating an equivalent copy, the changes of which have phonological values. From these simplified curves (obtained through IPO's close copy stylization or similar processes, e.g. the MOMEL/INTSINT one, see HIRST et al. 2000), it is possible to extract two types of data: either parameters defining continuous contours that span the segments of interest (syllable, prosodic word...), or discrete feature sets. Both approaches are primarily descriptive, but may be used to extract quantitative descriptions and comparison between prosodic performances.

### *Description based on contours*

Approaches of intonation have tried to describe the relations between  $F_0$  changes and time in terms of patterns of smooth curves; in many cases these approaches were linked to signal processing aims. The idea to use functions so to describe the non-linear changes of intonation movements along time was tested by Levitt & Rabiner (1971). They proposed to use orthogonal polynomial bases to that aim, and introduced two levels of analysis (based on short-term windows and on groups of continuously voiced syllables). Their method lacks an analysis of utterance-level contours, and also uses a strong assumption on time normalization (that fit their particular case). Olive (1975) analysed sentence-level intonation contours, with systematic variation of the utterance syntactic structure and length. In that respect, his work is similar to the constrained prosodic data gathered within the dialectological AMPER project (Contini et al. 2002). But Olive averages the  $F_0$  contours at the sentence level for different sentences with different phonemic content so to remove microprosodic effects. The averaged contours are then fitted by fourth-order orthogonal polynomials, and the models used for speech synthesis. Orthogonal polynomials were used for several studies targeting linguistic description; most of them target short linguistic elements that can be thought of as comparable across speakers and phrases, such as the syllable for tones (ANDRUSKI & COSTELLO 2004), the last prominence for sentence-final contours in English (LAI 2014). The coefficients of the polynomial bases are used to describe categories of intonation shapes and groups

productions into shape-coherent clusters. Similarly, Hadjipantelis et al. (2012) use a functional data analysis to fit smooth functions to the syllabic tones of Mandarin, and extract the most relevant shapes associated to each one.

Such approaches aim at summarizing a complex shape of time-varying  $F_0$  contour by a reduced number of parameters (in an information-theoretic approach); the Fujisaki Model has a similar background: fit a smooth curve, derived from a reduced set of parameters, to  $F_0$  variations (FUJISAKI 1983, 1988, 2004). Fujisaki approach is notably different from previous ones because the parameters are grounded in the physiological process of voice production, and thus each parameter has a specific interpretation (e.g. baseline  $F_0$  value, phrase or accent commands; see also Kochanski & Shih 2003 for an approach with similar physiological motivations). The Fujisaki model was implemented for production purposes; Mixdorff's implementation (2000) allows an automatic analysis of measured  $F_0$  variations in recorded datasets, so to automatically estimate the model's parameters from the raw measurements. It is then possible to use Fujisaki's model in a descriptive way, founding descriptions on sets of accent commands extracted from the analysis, and relating them to the production by speakers of voluntary melodic variations (in that respect, it does tie with IPO's philosophy of analysis, targeting meaningful intonation movements). Mixdorff's (2000) process starts by fitting the  $F_0$  values with a continuous smooth curve, based on MOMEL's spline fitting of raw  $F_0$  measures (HIRST & ESPESSER 1993). The MOMEL algorithm is similar to the functional fitting presented in the preceding paragraph, but it differs at least for the use of quadratic splines as fitting functions (MOMEL serves as a basis to calculate INTSINT's phonological features, cf. *infra*). From this smooth curve, the Fujisaki model estimates, *inter alia*, a set of accent commands (with their position in time, duration, and amplitude) that may be used as variables to describe the  $F_0$  contours, or in a similar way as feature sets (e.g. number of accented syllables, localization regarding the tonic syllable, etc.), so to describe prosodic variations – as Mixdorff & Pfitzinger (2005) have demonstrated.

### ***Description based on feature sets***

Describing prosodic changes in terms of features is typical of linguistic approaches to prosody – and follows the classical principle of feature sets for phonemic descriptions, as opposed to the mainly signal processing approach in the case of smooth contours. An early proposition for using prosodic features sets was made by Martin (1975, 1982), and improved in Martin (1987). He proposes

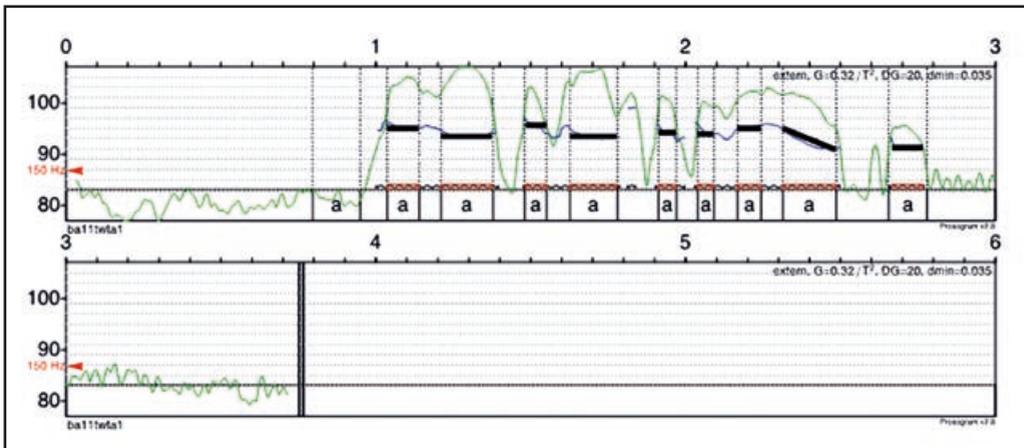
the use of feature matrices to describe prosodic markers. Another infamous approach to prosodic description, the ToBI system (SILVERMAN et al. 1992), and notably its application to Portuguese and Spanish varieties (FROTA & MORAES 2016; HUALDE & PRIETO 2015), may be seen as a series of features transcribing relevant phonological structures of prosody, or as categories the frequency of which vary across dialects (similarly to lexical approaches described in the introduction; CRUZ et al. 2017). In the same line, other systems of phonological transcription of prosody may serve the same purpose (e.g. INTSINT as in HIRST et al. 2000; Polytonia by MERTENS 2014).

Specifically targeting prosodic variation for dialectometric purposes, Contini & Profili (1989) build on Martin and IPO proposals to set up a 13-feature set describing variations of intonation and duration (see also Contini 1992 for a programmatic view). The features are established after an IPO-style stylisation process that allows straight lines representation of  $F_0$ . The features proposed for  $F_0$  changes on each syllable are  $\pm$ rising,  $\pm$ falling,  $\pm$ steep,  $\pm$ gentle,  $\pm$ wide,  $\pm$ narrow, plus two features indicating if the syllable is  $\pm$ above or  $\pm$ below the speaker's mean  $F_0$ . Five other features are attributed to one syllable, relatively to either the prosodic groups or the complete sentence: four features characterize the highest (vs. lowest)  $F_0$  point of the group or of the sentence, the strongest lengthening of the group or of the sentence; the last feature characterizes the  $\pm$ falling global  $F_0$  contour of the sentence. These binary features are applied to each vowel of sentences; that way, it is possible to create matrices describing a sentence's prosodic patterns, with features on lines, and vowels on columns. A column that gathers positive evaluations on most features is described as a "hot spot" (a "*point chaud*", in Contini's terms) of prosodic variation for this sentence.

Other approaches that extract sets of patterns or features from the  $F_0$  contours may be used to similar goals: the model of tonal perception proposed by d'Alessandro & Mertens (1995), and its implementation in Praat (BOERSMA & WEENINK 2018) as the Prosogram (MERTENS 2004) allow to automatically extract information on the pitch patterns composing sentences. Figure 1 shows an example of a sentence with  $F_0$  contour on vowels being stylized by the Prosogram (the example comes from the work of Rebollo Couto et al., 2017, that will be presented in more details latter). In that case, only the penultimate syllable (here the final tonic syllable) is stylized with a dynamic tone, all the other having contours simplified as a flat tone as their  $F_0$  variation is below the default threshold of perception for glissando used in the model of tonal perception (d'Alessandro & Mertens 1995). Such simplification of the raw  $F_0$  estimation (the

blue line) allows straightforwardly estimating features such as the  $F_0$  movement on each syllable (flat, raising, etc.), the amplitude of this movement, etc., so to construct features matrices similar to Contini & Profili (1989) proposition. The table presented at the bottom part of figure 1 is such a feature matrix. The eleven features are derived from Contini & Profili (1989) ones. The first three features describe the  $F_0$  movement on the vowel, if its flat or dynamic (raise, fall, raise-fall, etc.), the second the speed of these movements, and the third its amplitude; these features or not binary ones, contrary to Contini & Profili. Features 4 to 10 describe the vowel in relation with others: if its  $F_0$  is above or below the mean; if it is the highest of the prosodic word or of the utterance, if the syllable is the longest of the prosodic word or of the utterance, etc. The last one describes the direction of the sentence's slope. These features are presented in Rilliard (2014: p. 45ff); in figure 1, each column corresponds to (and is vertically aligned with) one vowel, and each line contains the description of one feature.

**Figure 1:** *Top* Prosogram stylization (thick black straight lines) of the  $F_0$  (blue line) of each of the 10 syllables of an assertive sentence from a female speaker of Rio de Janeiro (REBOLLO COUTO et al. 2017); each vowel is symbolized by an “a”, the first one not being produced in that case. The green line represents intensity. *Bottom* Feature matrix constructed from the Prosogram’s stylisation (see text).



Features	1	2	3	4	5	6	7	8	9	10
$F_0$ movement shape	0	0	0	0	0	0	0	0	-	0
$F_0$ movement speed	0	0	0	0	0	0	0	0	+	0
$F_0$ movement range	0	0	0	0	0	0	0	0	+	0
$F_0$ position ~ mean	0	0	0	0	0	0	0	0	+-	-
Highest (prosodic word)	0	0	-	+	-	+	0	0	+	-
Longest (prosodic word)	-	0	0	+	+	-	0	0	+	-
Most salient $F_0$ (prosodic word)	0	0	+	-	+	-	+	0	-	0
Highest (utterance)	0	0	0	0	0	0	0	0	+	-
Longest (utterance)	0	0	0	0	0	0	0	0	+	-
Most salient (utterance)	0	0	0	0	+	0	0	0	-	0
Sentence's slope	+	0	0	0	0	0	0	0	0	-

Grabe et al. (2007) proposed a detailed account on both types of approaches, based on feature sets and on smooth curves, so to evaluate the adequacy of the first approach (based on ToBI) to describe prosodic strategies across English urban varieties. The shapes of nuclear accents were analysed thanks to Legendre polynomials, with the  $F_0$  contours weighted by the signal's loudness and periodicity. Six out of the seven autosegmental-metrical label shapes were related to significantly different  $F_0$  contour shapes, with a potential limit due to data sparsity. This result enhances the coherence of both approaches, showing their coherent conclusions.

## QUANTITATIVE MEASUREMENT OF PROSODIC VARIATION

Once prosodic variation has been described in a systematic manner, whatever the model selected, the next task for a geolinguistic of prosody is to quantify its variation. Several propositions have been made for comparing other types of linguistic changes (reviewed in the introductory part) but can readily be applied to prosody; other proposals have been designed specifically for quantifying prosodic differences.

**Table 1:** reproduction of prosodic variations observed on sentences produced by one female and one male speaker from three cities, with two sentence modes (assertion: Ass., interrogation: Int.). Prosody is described in terms of autosegmental-metric labels (ToBI) and as AMPER stylized curves (red for assertion, blue for interrogation, 3  $F_0$  [measured in semitones] points per vowel, devoiced vowels marked by a low straight line).

City	Gender	Mode	ToBI	AMPER
Fortaleza	Female	Ass.	L+H* H+L*L%	
		Int.	L+H* H+H*L%	
	Male	Ass.	L+H* H+L*L%	
		Int.	L+H* L+H*L%	
Salvador	Female	Ass.	L+H* H+L*L%	
		Int.	L+H* L+H*L%	
	Male	Ass.	L+H* H+L*L%	
		Int.	L+H* L+H*H%	
Rio de Janeiro	Female	Ass.	H* H+L*L%	
		Int.	H* L+H*L%	
	Male	Ass.	H* H+L*L%	
		Int.	H* L+H*L%	

To apply different measurements of prosodic divergences as example of their application, I'll use the results of Rebollo Couto et al. (2017) who described prosody in three varieties of Brazilian Portuguese (Salvador, Fortaleza and Rio de Janeiro) following the AMPER methodology (Contini et al. 2002) and in terms of autosegmental-metric labels adapted to Portuguese (Frota & Moraes 2016). I'll take advantage of this versatility to show that several approaches are possible to get quantitative measurements of prosodic variations that may fit different research questions. In Rebollo Couto et al. (2017), the sentence “*O Renato gosta do Renato*” (Renato likes Renato) is presented for two speakers in each of these three cities, with assertive and interrogative modes. Table 1 summarizes its prosodic characteristics as seen within the autosegmental-metric framework (for the pre-nuclear and nuclear parts), and within the phonetic approach recommended by the AMPER project (in terms of the  $F_0$  contour). As an example, the feature matrix of one sentence is presented in figure 1.

### ***Divergence measures based on features***

Classical dialectometric approaches use different versions of edit distance metrics to compare sets of features; such distance measurement count the number of transformations (possibly deletion, insertion, substitution and transposition) required to convert one string of characters into another; many variants exist in the literature, allowing notably to weight the different operations (see Navarro 2001 for a technical survey, Heeringa 2004 for their use for dialectometry and Heeringa et al. 2006 for an evaluation of several forms of them). Such metrics apply directly to measures of prosodic features – being based on matrices or sets of labels.

Assertive sentences in table 1 all have the same nuclear ToBI pattern (H+L\*L%), differing only for the prenuclear part (L+H\* or H\*); their Damerau-Levenshtein distance (hereafter DL) equals two (that correspond to the missing “L+” at the beginning of Rio de Janeiro prenuclear accent). Thus for assertive mode, the Fortaleza and Salvador varieties have the same phonologic patterns (i.e. DL (Fortaleza, Salvador) = 0), while they both have a distance of 2 with the Rio de Janeiro variety. For interrogative sentences, the sum of differences between all pair of speakers from two different cities equals 10 for Rio de Janeiro / Fortaleza and Rio de Janeiro / Salvador, but 4 for Salvador / Fortaleza, as variations in the pre-nuclear part are systematic in Rio compared to the other two cities.

This string metric also applies to the feature matrices: the first feature (i.e. the pattern of  $F_0$  movement on each vowel) for interrogative sentences is respectively “0 0 + 0 0 0 0 0 + -”, “0 0 0 0 0 0 0 0 ++ 0”, “0 0 0 0 + 0 0 0 0+ -+” (with spaces separating features of each vowel) for the female and male speakers from Rio de Janeiro, and the female from Salvador; the two carioca speakers’ prosodic shapes diverge by a DL distance of 3, while each carioca speaker’s prosody has a DL distance of 4 with Salvador’s female production. The DL distance may be applied in a similar way to each feature (i.e. each pair of matrix’s lines), the sum of these distances across features giving an idea of the quantitative difference between prosodic patterns of two sentences.

Another possible approach would build on corpus based frequency analyses, looking at the frequency ratio between types of accents (or shapes of contours, etc.) occurring in each dialect under investigation. The approach defended for lexical analysis by Speelman & Geeraerts (2008) may here prove interesting: “profiles” of accentual patterns may be created, that list possible performances for a given prosodic function (e.g. assertion, interrogation), and extract the relative frequency of occurrence for each types of accentual patterns. In the current example, interrogation would possibly be actualized as “L+H\* H+H\*L%”, “L+H\* L+H\*L%”, “L+H\* L+H\*H%”, or “H\* L+H\*L%” under their ToBI description. A particular dialect will use these variants with a specific frequency (the Rio de Janeiro variety possibly focusing on one pattern only, while the two others cities may show a more diverse profile). The patterns of frequency distribution actually observed in a corpus for several prosodic functions may serve as a dependant variable for a statistical analysis that will compare such prosodic “profiles” across dialects (see Wieling & Nerbonne 2015 for a review). Details of the methodology are discussed, for lexical variation and its application to geolinguistic variation, in Speelman & Geeraerts (2008) – I don’t know any use of this concept for prosody, but it easily apply to, and may prove efficient as it comes with a robust methodology.

Application of regression modelling to the parameters of orthogonal polynomial decompositions allows evaluating the relative role of a set of factors (including geographic origin) to prosodic characteristic (see Kochansky et al. 2005; Grabe et al. 2007; Lai 2014). In such an approach, bounded contours are described by their shape components and their proximity may be evaluated and compared across geographic and/or social variation, showing the similarities between feature-based and some contour-based approaches. The parameters

extracted from an application of the Fujisaki model may be treated in a similar way so to analyse prosodic variation (MIXDORFF & PFITZINGER 2005).

### *Divergence measures based on contours*

When prosody is described as contours, other approaches are generally used to quantitatively compare them. An emblematic method was proposed and evaluated by Hermes (1998b), who shows the correlation between two  $F_0$  contours, weighted by the maximum amplitude of the subharmonic sumspectrum (a measure of the intensity in the speech signal that contributes to  $F_0$  perception; the measure was also applied elsewhere with signal intensity, or other loudness measurement), is the measure that best fit the perceived difference between two contours (HERMES 1998a). The measure is described in equation 1 (adapted from Hermes 1998b to the case of discrete  $F_0$  measures, as in d'Alessandro et al. 2011), where  $r$  is the weighted correlation of two vectors of  $n$   $F_0$  measures  $f_1$  and  $f_2$ ,  $w$  is a weighting vector of length  $n$ , and  $\mu_1$  and  $\mu_2$  are the respective means of  $f_1$  and  $f_2$ .

$$(1) \quad r_{f_1, f_2} = \frac{\sum_{i=1}^n w(i)(f_1(i) - \mu_1)(f_2(i) - \mu_2)}{\sqrt{\sum_{i=1}^n w(i)(f_1(i) - \mu_1)^2 \sum_{i=1}^n w(i)(f_2(i) - \mu_2)^2}}$$

This correlation measurement was applied to the comparison of prosodic performances for synthetic speech evaluation (HIRST et al. 1998; D'ALESSANDRO et al. 2011). The same measure was also used for the observation of prosodic divergences across dialects, in a seeding article on geoprosody (MOUTINHO et al. 2011). The measure is similar to others proposed for the same aim (see ROMANO et al. 2011 for an historic) but that propose unweighted measures (correlation or covariance) between two contours. Has shown by Hermes (1998b), considering the voicing strength adds much to the perceptual relevance of quantitative measurement, minimizing the influence of possibly ample  $F_0$  movements (e.g. linked to microprosodic effects or  $F_0$  rises after vowels, that may not be intended nor controlled) that have few or none functional relevance, as performed at very low intensity levels (and so are almost not perceived).

Some tools do implement the weighted correlation measures for dialectometric purposes; Martínez Calvo & Fernández Rei (2015) proposes an

implementation in the R software; Elvira-García et al. (2018) proposes an integrated tool based on this distance, and targeting AMPER-style data, but also data under a raw acoustic form. Their implementation of the distance allows varied forms of weight, interestingly one that possibly includes segmental duration as a weighting factor.

The weighted correlation measures a similarity between two sets of points; this raises several difficulties. First, the two  $F_0$  vectors to be compared shall have the same length, which requires a normalization so to deal with time differences between utterances (see Xu 1999 for an example of such time normalization), such linear normalizations can be contemplated only for cases of similar structure or length. Second, missing values raise difficulties when comparing two continua; such missing values are particularly frequent, but not systematic, for post-tonic final vowels in Brazilian Portuguese<sup>1</sup>. Third, its output is bounded between [-1; 1], so Hermes (1998b, p. 75) proposes to apply Fisher's Z-transform so to get values in the [0; +∞] range – the higher this measure, the more similar the two compared contours: it is thus (as mentioned by Hermes) a *similarity* measure that inversely correlate with *dissimilarity* measures between contours that have described up to now (e.g. string metric; on similarity and distance, see also Heeringa et al. 2006).

The numerical application of the weighted correlation gives the *similarities* (Z-transformed of the correlations) reported in table 2, for the AMPER-stylization reported in table 1. All pairs of sentences with the same modality have been compared, which makes three types of pairs: same-speaker pairs, same-city pairs (two speakers from the same city), and different-city pairs. Means of these measures are reported, aggregating all pairs from one city together (same speaker or not), separately for each modality. The mean similarity for assertive sentences is 1.06, and 0.59 for interrogative ones: there is much more cross-dialectal variation in interrogative performances, as the results obtained from ToBI labels had already shown.

---

<sup>1</sup> Solutions exist, such as interpolation, but have to be adapted specifically to each situation.

**Table 2:** mean of the weighted correlation's Z-transforms between each pair of sentences presented in table 1 (same modality) between speakers from the three cities (Rio de Janeiro, Salvador and Fortaleza). See text for details.

Mode	City	Rio de Janeiro	Salvador	Fortaleza
Assertion	Rio de J.	1.26	1.10	0.87
	Salvador		1.58	0.83
	Fortaleza			0.85
Interrogation	Rio	1.05	0.36	0.51
	Salvador		1.08	0.19
	Fortaleza			0.61

The correlation measure between speakers of the same city gives information on the coherence of the speakers in terms of prosodic strategy, and thus on variability of prosodic patterns there. For example, the two speakers from Fortaleza show low similarity rating, compared to what is observed in other cities (especially for interrogatives: 0.61 vs. values superior to 1). Comparing intra-individual variation (i.e. coherence of one speaker to reproduce similar pattern across repetitions) and inter-speaker measures (for the same city), is possible to propose inferences on the representativeness of measured contours as dialect-specific, while inter-city measurements are indicative of variation between varieties<sup>2</sup> (for such an approach, even if based on another measure, see Grabe et al. 2007).

In our example, interrogative productions from Salvador are more dissimilar to those of other cities. Note that the results found here are not completely coherent with what was calculated on the basis of the ToBI labels (*supra*). The large difference between Rio de Janeiro and the two north-eastern cities observed with ToBI labels is mostly related to the different encoding of the prenuclear part, while this part (also taken into account in the weighted correlation) sounds much less important in its magnitude than the discrepant changes in the nucleus. One may also apply the LD metric on nuclear accent only, but the greater difference between Fortaleza and Salvador (compared to Rio de Janeiro) would remain, as one speaker of both cities diverge from the phonological pattern described for Rio de Janeiro (Moraes 1998), but each in a specific way. Knowing how these

<sup>2</sup> The example data presented here are not sufficient for proposing a statistical approach, but examples exist in the literature: see Grabe et al. 2007 or Wieling et al. 2014.

variations are perceived, which variant is perceived as most distant from the others, would allow tuning quantitative metrics so they may more accurately reflect perception (by weighting specific parameters or operation: a string metric may weight differently each operation, for example by putting more load to deletion of a character than on its substitution). It is also important to note the example here is based on a ridiculously small subset; using more data allows a more robust analysis, and may reconcile the different measures.

As it was said earlier, measures of correlation (or similar ones) have problems linked with variations in the duration of spoken utterances to be compared. The most common solution consists in time normalization, as described in Xu (1999), and that acts at a phonemic level. Meanwhile, this approach can hardly solve duration differences linked with variation in structure (different sets of phonemes, even with equivalent number of syllables or other targeted structures). Rilliard et al. (2011) did propose the use of a non-linear adaptation of prosodic contour before applying the correlation measure. The non-linear process is based on the application of a dynamic-time-warping (DTW) algorithm (as in Jouanelle et al. 1981) to vectors of prosodic characteristics that includes the intonation contours but also duration and intensity patterns. The results show non-linear alignment allows reaching higher similarity measurements for shape-similar, but time varied contours. The DTW algorithm show strong similarities with string metrics, in an approach applied to signal processing: both target an optimal path to compare or transform pairs of vectors according to their similarities; DTW works on numeric values and string metrics on character. Both of these approaches aim at finding the relation between two sequences that could be represented as acoustic signals or as strings. Heeringa & Gooskens (2003) application of the Levenshtein distance to acoustic data is an excellent demonstration of this convergence.

## AGGREGATION OF DISTANCES

Once a metric has been applied to the comparison of prosodic performances within and across speakers and dialectal areas, the dialectometric process consists in an agglomeration of these measurements so to reach representation of the variation at hand before descriptive and interpretative works. This part is not specific to prosody: similar methods can be applied to prosodic divergences that were already used since Goebel (1993, 2003, 2006). It basically consists in aggregating the output of objective metrics into a “distance matrix” (eventually converting similarity measures into divergence ones); this matrix is a symmetric

table the cells of which present the metric's values that separate two *individuals* (individuals may be speakers, sentences, localities etc.): the diagonal presents the internal divergence within each individual (if any, it may be zeros), while the other cells represent between-individuals divergences.

This matrix, which may be very large, is then subjected to a multidimensional analysis (e.g. HUSSON et al. 2017 for a theoretical and practical account). This analysis will spread all individuals on a few abstract dimension that best structure the cloud of individuals. The Multidimensional Scaling (MDS) method is very often used, as it allows forcing the spread of individuals on the bidimensional output (i.e. a plane) that best represent the observed variation, and that easily fits on a map. Other favourite algorithms used to deal with distance matrices are clustering ones – and typically hierarchical clustering algorithm, that produces dendrogram representation of the individuals by iteratively grouping (or splitting) them in wider (or smaller) groups (see Husson et al. 2017, chapter 4, for details).

Such outputs are ideally used by a geographic information system (GIS) so to plot the result on a map – one frequent requirement for dialectology. A few GIS specialized in the representation of dialectological data: GabMap is a prominent one (many other flavours exist as a search of “GIS dialectology” in a search engine may show) that also comes with an R interface (<http://www.let.rug.nl/~kleiweg/L04/>) and allows geographic representation of the variation captured in a metric. The type of dialectological information to be represented varies: distance from a reference point (as a shade of colour), boundaries between variants or links within variants, or aggregated patches of clustered variants (e.g. Goebel 2006, Nerbonne 2009).

Another approach, much typical of sociolinguistics, is based on regression analysis, and may allow the observation of several levels of variations – social and geographic. The notion of distance is not necessarily summarized in a matrix, but still can be represented on a map, as the work presented in Wieling et al. (2014) may demonstrate. Wieling & Nerbonne (2015) proposes an excellent review of methods and approaches with lists of existing tools.

## DISCUSSION & CONCLUSIONS

This article proposes a review of methods used to measure objective divergences between prosodic performances – typically in relation with dialectological variation. We have seen the prosodic aspect of speech is still under-resourced

in the field of dialectology and sociolinguistic. Both field proposes some solutions, which may be seen as different, but do recently show their possibility to converge, and their potential application to other type of data. The main challenge (as prosody is concerned) certainly remains reliable objective measures of prosodic variation; other tools and methodologies (statistical aggregations and analysis, multidimensional analyses, and representation tools) have been already developed for the other levels of linguistic description and are more or less readily applicable to the case of prosody.

Propositions for such metrics have been reviewed, that show representation of prosody is the basis to set up divergence measurements. As many representation of prosody are linked to theoretical models (e.g. the ToBI annotation scheme to the phonological autosegmental-metric theory), the choice of one solution will generally not be neutral, and may be dictated by the special interest of a particular research: the AMPER project for example as a strong phonetic focus that rely on its aims to observe variation as a first step to describe it, and then having the capacity to categorize it. Grabe et al. (2007) have shown approaches based on very different materials may show a strong convergence; thus a readily applicable choice is certainly the better solution for a given research question, that will depend on the researcher familiarity with the various aspects of such approaches, and access to / familiarity with the related tools. The presentation of different measurements based on the same example dataset shows that there are similarities to be found between approaches, but also potentially differences linked to the importance placed on different aspects of the prosodic changes.

One such difference is related to the set of parameters used to describe prosody, and notably on the use of duration related measurements. Most models of prosody focus on intonation, which has certainly a prime importance – but rhythm and lengthening also carry critical information on the linguistic message (see Moraes 2008 for a discussion on some lengthening effects). The matrices of features advocated by Contini (1992) in his proposition of a geoprosodic approach to dialectal variation may include such durational level of information; Mixdorff & Pfitzinger (2005) proposes another possible solution based on Pfitzinger's (1998) proposal of a “perceptual local speech rate”, that gives a continuous measure of rhythm. This approach shares similarities with the models of duration proposed by Campbell (1993) and Barbosa (2007). Such continuous representations of duration may be taken into account together with  $F_0$  by approaches based on functional modelling – as well as other parameters as loudness or aperiodicity – as was done in Kochanski & Shih (2003) and Kochanski et al. (2005).

So to get more reliable tools, problems still need to be further investigated. The main one certainly is linked with the relations between the acoustic dimension of speech and its perceptual interpretation that is not done in a unique and systematic way for a given set of audio parameters, and neither is similar across language varieties. Complex challenges include setting up objective measures able to take into account pragmatic variables so to deal with changes linked to expressive speech (because of e.g. hierarchical relations between interlocutors; vocal effort control due to situation: noise, distance, need for discretion; expected behaviour in relation to social norms). Another challenge is related to the limited set of prosodic patterns compared to the complexity of pragmatic interpretation: prosodic meaning consists in a small set in comparison to pragmatic uses of language, but there is still few knowledge on the composition of this set (see Mixdorff et al. 2017 for a discussion).

A difficulty linked to these challenge is related to the complexity of some prosodic modelling, that involves complex statistical approaches, and that are not necessarily proposed within user-friendly computer program. Comparatively, the success of the varbrul analysis (CEDERGREN & SANKOFF 1974) in sociolinguistics is certainly due to the efficient computer programs that allow efficient applications of an otherwise quite complex mixed effects logistic regression. Some tools are available to apply dialectometric measures (see Wieling & Nerbonne 2015 for a review) and a few tools specialized in measuring prosodic divergences (notably Martínez Calvo & Fernández Rei 2015; Elvira-García et al. 2018), but there are still spaces for much work in that direction.

Finally, there is also questions linked with the varying sources of prosodic variation and its representation: if dialectology focuses mostly on the diatopic aspect, there is certainly much differences linked to the diastratic composition of the society, especially in urban areas and in cities where important difference in education are found within the population. Geographic changes are efficiently represented through techniques pioneered by Goebel (see the introduction), but they are not necessarily adapted to take into account important differences in the special density of survey points, and especially cases where variation is more social than geographic – a fact that is more and more common with population mobility and important urban concentrations. Reflexion and interdisciplinary collaborations (notably with geographers and computer vision scientists) are needed so to propose and adapt solutions to concrete cases of geoprosody.

## ACKNOWLEDGMENTS

The author wants to thank the invaluable support of UFRJ in the completion of this article; I'm also indebted to Leticia Rebollo Couto to let me use her data as examples for this article, and Jussara Abraçado for discussions on onomasiological profiles.

## REFERENCES

D'ALESSANDRO, Christophe; MERTENS, Piet. Automatic pitch contour stylization using a model of tonal perception. *Computer Speech and Language*, v. 9, n. 3, p. 257, 1995.

D'ALESSANDRO, Christophe; RILLIARD, Albert; LE BEUX, Sylvain. Chiro-nomic stylization of intonation. *The Journal of the Acoustical Society of America*, v. 129, n. 3, p. 1594-1604, 2011.

ANDRUSKI, J. E.; COSTELLO, J. Using polynomial equations to model pitch contour shape in lexical tones: an example from Green Mong. *Journal of the International Phonetic Association*, v. 34, n. 02, p. 125-140, 2004.

BAAYEN, R. Harald. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge University Press, 2008.

BARBBOSA, Plinio. From syntax to acoustic duration: A dynamical model of speech rhythm production. *Speech Communication*, v. 49, n. 9, p. 725-742, 2007.

BAYLEY, Robert. The quantitative paradigm. In Chambers, J. K., Peter Trudgill and Natalie Schilling-Estes (eds) *The handbook of language variation and change*, p. 117-141, 2003.

VAN BEZOOIJEN, Renée; GOOSKENS, Charlotte. Identification of language varieties: The contribution of different linguistic levels. *Journal of language and social psychology*, v. 18, n. 1, p. 31-48, 1999.

BOERSMA, Paul; WEENINK, David. Praat: doing phonetics by computer [Computer program]. Version 6.0.35. Online: <http://www.praat.org/>. Retrieved in 2018.

CAMPBELL, Nick. Automatic detection of prosodic boundaries in speech. *Speech Communication*, v. 13, n. 3, p. 343–354, 1993.

CEDERGREN, Henrietta J.; SANKOFF, David. Variable rules: Performance as a statistical reflection of competence. *Language*, v. 50, n. 2, p. 333-355, 1974.

CHAMBERS, Jack K.; TRUDGILL, Peter. *Dialectology* (second edition). Cambridge University Press, 2004.

CONTINI, Michel. Vers une géoprosodie. In: *Proceedings of the Nazioarteko Dialektologia Biltzarra, IKER 7, Euskaltzaindia*. Bilbao: Real Academia de la Lengua Vasca, p. 83-109, 1992.

CONTINI, Michel; PROFILI, Olga. L'intonation de l'italien régional – un modèle de description par traits. In BOTHOREL, A. et al. (Eds.), *Mélanges de phonétique générale et expérimentale offerts à Péla Simon*, Strasbourg: Publications de l'Institut de Phonétique de Strasbourg, p. 855-870, 1989.

CONTINI, Michel; LAI, Jean-Pierre; ROMANO, Antonio; ROULLET, Stefania; MOUTINHO, Lurdes de Castro; COIMBRA, Rosa Lídia; BENDIHA, Urbana Pereira; RUIVO, Suzana Secca. Un projet d'atlas multimédia prosodique de l'espace roman. In: *Proceedings of the International Conference of Speech Prosody*, p. 227-230, 2002.

CRUZ, M.; OLIVEIRA, P.; PALMA, P.; NETO, B.; FROTA, S. Building a prosodic profile of European Portuguese varieties: The challenge of mapping intonation and rhythm. In BARBOSA, P.; PAIVA, M. C.; RODRIGUES, C. (Eds.), *Studies on Variation in Portuguese*. Amsterdam: John Benjamins, p. 81-110, 2017.

ELVIRA-GARCÍA, Wendy; BALOCCO, Simone; ROSEANO, Paolo; FERNÁNDEZ-PLANAS, Ana Ma. ProDis: A dialectometric tool for acoustic prosodic data. *Speech Communication*, 2018, vol. 97, p. 9-18.

FROTA, Sónia; MORAES, João Antônio. Intonation in european and brazilian portuguese. In: Wetzels, W.L; Menuzzi, S.; Costa, J. (Eds.), *Handbook of Portuguese Linguistics*. 1ed. Hoboken: Wiley-Blackwell, p. 141-166, 2016.

FUJISAKI, Hiroya. Dynamic characteristics of voice fundamental frequency in speech and singing. In MACNEILAGE, P. (Ed), *The production of speech*. New York: Springer, p. 39–55, 1983.

FUJISAKI, Hiroya. A note on the physiological and physical basis for the phrase and accent components in the voice fundamental frequency contour. In FUJIMURA, O. (Ed.), *Vocal fold physiology: voice production, mechanisms and functions*. New York: Raven, p. 347–355, 1988.

FUJISAKI, Hiroya. Information, prosody, and modelling - with emphasis on tonal features of speech. In: *Proceedings of the International Conference on Speech Prosody*, Nara, Japan, p. 1-10, 2004.

GOEBL, Hans. Dialectometry: a short overview of the principles and practice of quantitative classification of linguistic atlas data. In: Reinhard Köhler; Burghard B. Rieger (Org.) *Contributions to quantitative linguistics*. Springer, Dordrecht, p. 277-315, 1993.

GOEBL, Hans. Regards dialectométriques sur les données de l'Atlas linguistique de la France (ALF): Relations quantitatives et structures de profondeur. *Estudis Romànics*, p. 60-117, 2003.

GOEBL, Hans. Recent advances in Salzburg dialectometry. *Literary and linguistic computing*, v. 21, n. 4, p. 411-435, 2006.

GOOSKENS, Charlotte Stenkilke. On the role of prosodic and verbal information in the perception of Dutch and English language varieties. PhD Thesis, Katholieke Universiteit Nijmegen, 1997.

GOOSKENS, Charlotte; HEERINGA, Wilbert. Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language variation and change*, v. 16, n. 3, p. 189-207, 2004.

GOOSKENS, Charlotte; HEERINGA, Wilbert. The relative contribution of pronunciational, lexical, and prosodic differences to the perceived distances between Norwegian dialects. *Literary and Linguistic Computing*, v. 21, n. 4, p. 477-492, 2006.

GOOSKENS, Charlotte. Non-linguists' judgments of linguistic distances between dialects. *Dialectologia: revista electrònica*, n. 9, p. 27-51, 2012.

GOOSKENS, Charlotte. Experimental methods for measuring intelligibility of closely related language varieties. *The Oxford handbook of sociolinguistics*, p. 195-213, 2013.

GOOSKENS, Charlotte. Methods for measuring intelligibility of closely related language varieties. In: BAYLEY, Robert; CAMERON, Richard; LUCAS, Ceil (eds.). *Handbook of sociolinguistics*. Oxford University Press, p. 195-213, 2013.

GRABE, Esther; KOCHANSKI, Greg; COLEMAN, John. Connecting intonation labels to mathematical descriptions of fundamental frequency. *Language & Speech*, v. 50, n. 3, p. 281-310, 2007.

HADJIPANTELOS, Pantelis Z.; ASTON, John AD; EVANS, Jonathan P. Characterizing fundamental frequency in Mandarin: A functional principal component approach utilizing mixed effect models. *The Journal of the Acoustical Society of America*, v. 131, n. 6, p. 4651-4664, 2012.

'T HART, Johan; COLLIER, René; COHEN, Antonie. A perceptual study of intonation: an experimental-phonetic approach to speech melody. Cambridge: Cambridge University Press, 1990.

HEERINGA, Wilbert Jan. Measuring dialect pronunciation differences using Levenshtein distance. PhD, University Library Groningen. 2004.

HEERINGA, Wilbert; NERBONNE, John. Dialect areas and dialect continua. *Language Variation and Change*, v. 13, n. 3, p. 375-400, 2001.

HEERINGA, Wilbert; GOOSKENS, Charlotte. Norwegian dialects examined perceptually and acoustically. *Computers and the Humanities*, v. 37, n. 3, p. 293-315, 2003.

HEERINGA, Wilbert; KLEIWEG, Peter; GOOSKENS, Charlotte; NERBONNE, John. Evaluation of string distance algorithms for dialectology. In: *Proceedings of the workshop on linguistic distances*. Association for Computational Linguistics, p. 51-62, 2006.

HEERINGA, Wilbert; NERBONNE, John. Dialectometry (Chap. 33). In HINSKENS, Frans; TAELEDAMAN, Johan (eds.): *Language and Space. An International Handbook of Linguistic Variation, Volume III: Dutch*. (Series: *Handbook of Linguistics and Communication Science (HSK)*). Berlin and New York: Walter de Gruyter, p. 624-646, 2013.

HERMES, Dik J. Auditory and visual similarity of pitch contours. *Journal of Speech, Language, and Hearing Research*, vol. 41, no 1, p. 63-72, 1998a.

HERMES, Dik J. Measuring the perceptual similarity of pitch contours. *Journal of Speech, Language, and Hearing Research*, vol. 41, no 1, p. 73-82, 1998b.

HIRST, Daniel; ESPESSER, Robert. Automatic modelling of fundamental frequency using a quadratic spline function. *Travaux de l'Institut de Phonétique d'Aix*, v. 15, p. 75-85, 1993.

HIRST, Daniel; RILLIARD, Albert; AUBERGÉ, Véronique. Comparison of subjective evaluation and an objective evaluation metric for prosody in text-to-speech synthesis. In: Proceedings of the Third ESCA/COCOSDA Workshop (ETRW) on Speech Synthesis, p. 1-4, 1998.

HIRST, Daniel; DI CRISTO, Albert; ESPESSER, Robert. Levels of representation and levels of analysis for the description of intonation systems. In: Merle Home (Ed.) *Prosody, theory and experiment: studies presented to Gösta Bruce*. Springer, Dordrecht, p. 51-87, 2000.

HUALDE, José Ignacio; PRIETO, Pilar. Intonational variation in Spanish: European and American varieties. In: FROTA, Sónia; PRIETO, Pilar (Eds.), *Intonation in Romance*. Oxford: Oxford University Press, p. 350-391, 2015.

HUSSON, François; LÊ, Sébastien; PAGÈS, Jérôme. *Exploratory multivariate analysis by example using R*. Chapman and Hall/CRC, 2017.

INOUE, Fumio. Computational dialectology (2). *Area and Culture Studies*, v. 53, p. 115-134, 1996.

INOUE, Fumio. A century of language change in progress. New dialect in Tsuruoka, Japan. *Dialectologia: revista electrònica*, v. 17, p. 71-89, 2016.

JOUANNELLE, G.; BOË, L. J.; BELLET, G. Variation intra et inter locuteur de F0: procédures de normalisation – étude préliminaire. In: *Proceeding of the 12es JEP*, Montréal, 246–262, 1981.

KOCHANSKI, Greg; SHIH, Chilin. Prosody modeling with soft templates. *Speech Communication*, v. 39, p. 311-352, 2003.

KOCHANSKI, Greg, GRABE, Esther, COLEMAN, John, et al. Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, v. 118, n. 2, p. 1038-1054, 2005.

KUIPER, Lawrence. Parisian Perceptions of Regional French. In PRESTON, Dennis R. (Ed.), *Handbook of perceptual dialectology*, v. 1, p. 243-262, 1999.

LAI, C. Interpreting Final Rises: Task and Role Factors. In: *Proceedings of Speech Prosody*, Dublin, Ireland, p. 520-524, 2014.

LEVITT, Harry; RABINER, Lawrence R. Analysis of fundamental frequency contours in speech. *The Journal of the Acoustical Society of America*, v. 49, n. 2-2, p. 569-582, 1971.

MARTIN, Philippe. Analyse phonologique de la phrase française. *Linguistics*, v. 13, n. 146, p. 35-68, 1975.

MARTIN, Philippe. Phonetic realisations of prosodic contours in French. *Speech communication*, v. 1, n. 3-4, p. 283-294, 1982.

MARTIN, Philippe. Prosodic and rhythmic structures in French. *Linguistics*, v. 25, n. 5, p. 925-950, 1987.

MARTIN, Philippe. Structures prosodiques des langues romanes. In: *Actes de la conférence conjointe JEP-TALN-RECITAL 2016*, volume 1: JEP, p. 678-686, 2016.

MARTÍNEZ CALVO, Adela; FERNÁNDEZ REI, Elisa. Unha ferramenta informática para a análise dialectométrica da prosodia. *Estudios de fonética experimental*, v. 24, p. 289-303, 2015.

MASE, Yashio. On Dialect Consciousness – Dialect Characteristics Given by Speakers. In PRESTON, Dennis R. (Ed.), *Handbook of perceptual dialectology*, v. 1, p. 101-113, 1999.

MERTENS, Piet. The prosogram: Semi-automatic transcription of prosody based on a tonal perception model. In: *Proceedings of the International Conference on Speech Prosody*. 2004.

MERTENS, Piet. Polytonia: a system for the automatic transcription of tonal aspects in speech corpora. *Journal of Speech Sciences*, v. 4, n. 2, p. 17-57, 2014.

MIXDORFF, Hansjörg. A novel approach to the fully automatic extraction of Fujisaki model parameters. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 00CH37100)*. IEEE. p. 1281-1284, 2000.

MIXDORFF, Hansjörg; PFITZINGER, Hartmut R. Analysing fundamental frequency contours and local speech rate in map task dialogs. *Speech Communication*, v. 46, n. 3-4, p. 310-325, 2005.

MIXDORFF, Hansjörg, HÖNEMANN, Angelika, RILLIARD, Albert, LEE, Tan; MA, Matthew K. A. Audio-visual expressions of attitude: How many different attitudes can perceivers decode?. *Speech Communication*, v. 95, p. 114-126, 2017.

MORAES, J. A. de. Intonation in Brazilian Portuguese. In: HIRST, Daniel; DI CRISTO, Albert. (Org.). *Intonation Systems: a Survey of Twenty Languages*. Cambridge: Cambridge University Press, p. 179-194, 1998.

MORAES, J. A. de. The pitch accents in Brazilian Portuguese: analysis by synthesis. In: *Proceedings of the International Conference of Speech Prosody*, v. 4, p. 389-397, 2008.

MOUTINHO, L. de Castro; COIMBRA, R. L.; RILLIARD, A.; ROMANO, A. Mesure de la variation prosodique diatopique en portugais européen. *Estudios de fonética experimental*, v. 20, p. 33-55, 2011.

NAVARRO, Gonzalo. A guided tour to approximate string matching. *ACM computing surveys (CSUR)*, v. 33, n. 1, p. 31-88, 2001.

NERBONNE, John. Data-driven dialectology. *Language and Linguistics Compass*, v. 3, n. 1, p. 175-198, 2009.

NERBONNE, John; COLEN, Rinke; GOOSKENS, Charlotte; KLEIWEG, Peter; LEINONEN, Therese. Gabmap-a web application for dialectology. *Dialectologia: revista electronica*, Special issue II, p. 65-89, 2011.

OLIVE, Joseph P. Fundamental frequency rules for the synthesis of simple declarative English sentences. *The Journal of the Acoustical Society of America*, 1975, v. 57, n. 2, p. 476-482.

PEIRSMAN, Yves; GEERAERTS, Dirk; SPEELMAN, Dirk. The automatic identification of lexical variation between language varieties. *Natural Language Engineering*, v. 16, n. 4, p. 469-491, 2010.

PFITZINGER, Hartmut R. Local speech rate as a combination of syllable and phone rate. In: *Proceedings of the fifth international conference on spoken language processing*, 1998.

REBOLLO COUTO, Leticia; GOMEZ DA SILVA, Carolina; DA SILVA MIRANDA, Luma. Prosódia de enunciados declarativos e interrogativos totais nas variedades de Salvador, Fortaleza e Rio de Janeiro. *Revista de Estudos da Linguagem*, v. 25, n. 3, p. 1105-1142, 2017.

RILLIARD, Albert. *Prosodie et Interaction Homme-Machine : Étude de la variation démarcative, diatopique, diachronique & expressive*. Mémoire d'Habilitation à Diriger des Recherches, Université Paris Sud, France, 2014.

RILLIARD, Albert; ALLAUZEN, Alexandre; BOULA DE MAREÛIL, Philippe. Using dynamic time warping to compute prosodic similarity measures. In: *Proceedings of the twelfth Annual Conference of the International Speech Communication Association (INTERSPEECH)*, p. 2021-2024, 2011.

ROMANO, Antonio; CONTINI, Michel; LAI, Jean Pierre; RILLIARD, Albert. Distancias prosódicas entre variedades románicas en el marco del proyecto AMPER. *Revista Internacional de Lingüística Iberoamericana*, v. 9, n. 1, p. 13-25, 2011.

ROMNEY, A. Kimball; MOORE, Carmella C., BATCHELDER, William H., & HSIA, Ti-Lien. Statistical methods for characterizing similarities and differences between semantic structures. *Proceedings of the National Academy of Sciences*, v. 97, n. 1, p. 518-523, 2000.

SÉGUY, Jean. La relation entre la distance spatiale et la distance lexicale. *Revue de linguistique romane*, v. 35, p. 335-357, 1971.

SILVERMAN, Kim; BECKMAN, Mary; PITRELLI, John; OSTENDORF, Mari; WIGHTMAN, Colin; PRICE, Patti; PIERREHUMBERT, Janet; HIRSCHBERG, Julia. ToBI: A standard for labeling English prosody. In: *Second International Conference on Spoken Language Processing (ICSLP)*, Banff, Canada, p. 867-870, 1992.

SPEELMAN, Dirk; GEERAERTS, Dirk. The role of concept characteristics in lexical dialectometry. *International Journal of Humanities and Arts Computing*, v. 2, n. 1-2, p. 221-242, 2008.

WEINER, E. Judith; LABOV, William. Constraints on the agentless passive. *Journal of linguistics*, v. 19, n. 1, p. 29-58, 1983.

WIELING, Martijn; MONTEMAGNI, Simonetta; NERBONNE John; BAAYEN, R. Harald. Lexical differences between Tuscan dialects and standard Italian: accounting for geographic and socio-demographic variation using generalized additive mixed modeling. *Language*, v. 90, n. 3, p. 669 – 692, 2014

WIELING, Martijn; NERBONNE, John. Advances in dialectometry. *Annual Review of Linguistics*, v. 1, n. 1, p. 243-264, 2015.

XU, Yi. Effects of tone and focus on the formation and alignment of F0 contours. *Journal of phonetics*, v. 27, n. 1, p. 55-105, 1999.



# SOCIOLINGUÍSTICA VARIACIONISTA E GRAMÁTICA DE CONSTRUÇÕES OS DESAFIOS E AS PERSPECTIVAS DE COMPATIBILIZAÇÃO

*Marcia dos Santos Machado Vieira (UFRJ)*  
*Marcos Luiz Wiedemer (UERJ)*

### RESUMO

Neste capítulo, defendemos um tratamento do fenômeno de variação no âmbito da Gramática de Construções do Português a partir da compatibilização de orientações e conceitos da Sociolinguística e da Gramática de Construções. No intuito de construir uma heurística socioconstrucionista, lançamos mão de três constructos teórico-metodológicos: (i) variação por aloconstruções/metaconstrução; (ii) variação por semelhança simbólica; e (iii) variação por paradigma/padrão discursivo. Sinalizamos um encaminhamento de trabalho científico que – ao centrar-se na experiência de uso ou de processamento, percepção ou avaliação subjetiva, bem como considerar uma rede de padrões construcionais configurada por introspecção e empiria – desenvolva generalizações estatísticas sobre os *links* associativos entre fatores de atributos das faces forma e função de construções postas/percebidas no uso em relação de variação por similaridade/sinonímia. Em linhas gerais, nossa contribuição já vislumbra um caminho que articule análise multivariada e *collostructional analysis* para lidar com conexões entre padrões construcionais variantes e com padrões de atração entre colexemas a um mesmo *slot* construcional ou padrões de coerção de lexemas a um *slot* construcional.

## INTRODUÇÃO

Retomamos a discussão seminal feita por Machado Vieira e Wiedemer (no prelo), e exposta no encontro do GT de Sociolinguística da ANPOLL em 2017, sobre o tratamento da variação no modelo da Gramática de Construções, um referencial teórico de gramática em Linguística Funcional-Cognitiva que se baseia na experiência.<sup>1</sup> Nessa ocasião, lançamos mão de três constructos teórico-metodológicos: (i) variação por aloconstruções/metaconstrução; (ii) variação por semelhança simbólica; e (iii) variação por paradigma/padrão discursivo.

Neste capítulo, concentramo-nos na discussão da “variação construcional” e de como abordagens construcionistas centradas no uso podem lidar com esse fenômeno. Em outras palavras, focalizamos a configuração de variável linguística numa perspectiva de articulação entre Sociolinguística Variacionista e Gramática de Construções, num enfoque que temos intitulado de *socioconstrucionista* (cf. WIEDEMER; MACHADO VIEIRA, 2018a e WIEDEMER; MACHADO VIEIRA, 2018b). Para tal finalidade, aqui, problematizamos duas dentre as possibilidades de variação: (i) variação por analogia/alinhamento de construções independentes e (ii) variação por compatibilização de (co)lexemas a *slots* de construção. Procuramos reunir procedimentos analíticos para lidar com esses tipos de variação.

Inicialmente, é importante apontar que concebemos a língua como um sistema cognitivo e social de construções (unidades simbólicas que pareiam atributos relativos à forma linguística – prosódica, fonético-fonológica, morfológica, sintática, lexical – a atributos relativos à funcionalidade/significação semântica, discursiva, pragmática, social e cognitiva). Esse sistema licencia<sup>2</sup> a linguagem nas experiências de uso e de processamento (mental/emocional) da língua e é por estas moldado. E, então, se sujeita a processos cognitivos gerais como *pensamento analógico* (processo de combinação de aspectos/atributos de significação/função e forma) que podem ensejar ou não analogização (mecanismo de mudança que faz emergir uma correspondência até então inexistente) e *parsing* (processo de análise em meio a uma complexidade de dados/*inputs* que potencializa, como *output*, (i) reconhecimento de unidades simbólicas ou (ii) reconfiguração de relações, recombinação diferente da análise corrente)

---

<sup>1</sup> E, ao encaminharmos essa relação, estamos, também, respondendo a questões propiciadas pela análise heurística de Machado Vieira (2016), conforme discussão final.

<sup>2</sup> Entende-se que o modelo de gramática na mente do falante se (re)configura com base no uso e também opera licenciando tokens/ usos, numa relação em dois sentidos (bottom-up e top-down).

que pode ensejar ou não neanálise (mecanismo de mudança que faz emergir uma nova estrutura). Mudança implica a etapa de espraiamento social e convencionalização de uma inovação linguística (individual) numa comunidade.

Concebemos a língua como um inventário de construções que envolvem estabilidade (relativa) e instabilidade (heterogeneidade e dinamicidade). Esta não se circunscreve ao fenômeno de mudança<sup>3</sup>, ou seja, a alteração ao longo do tempo (aparente ou real), como muitas pesquisas construcionistas (que normalmente, quando consideram variação, a relacionam à polissemia ou à degeneração (VAN DE VELDE, 2014)) podem levar a pensar. Se construções são unidades simbólicas regulares que licenciam os enunciados que usamos e processamos, mas que também estão sujeitas a instabilidade – em razão de fatores internos às construções, externos à atualização destas em constructos/ usos e a fatores externos e internos aos falantes em situações reais de experiência linguística e/ou de processamento linguístico –, consequentemente é central o papel da variação nas generalizações na Gramática de Construções de uma língua. Naturalmente, esse papel pode ser, além de central, periférico, a depender, na verdade, do tipo de problema para o qual se incline a proposta de pesquisa (variação; ou variação-mudança, mudança)<sup>4</sup>.

Tendo em vista tais concepções, importam generalizações tanto sobre usos situados social, histórica, cultural, cognitiva<sup>5</sup>, semântica, discursiva, pragmática

---

<sup>3</sup> Segundo Traugott e Trousdale (2013), esse processo ou é *mudança construcional* (alteração num atributo da face forma ou num atributo da face função/significação de uma construção, sem desencadear um novo nó construcional na rede (de construções) que constitui o sistema linguístico) ou é *construcionalização* (alteração tanto na forma quanto na função/significação que gera um novo nó/padrão construcional – forma nova pareada a função/significação nova). Na perspectiva construcionista de mudança, conforme Traugott e Trousdale (2013), **inovações** são características de conhecimentos individuais e se manifestam nas redes construcionais individuais, enquanto **mudanças** decorrem de um processo de convencionalização numa comunidade e, então, são compartilhadas. Em outras palavras, mudanças na rede construcional comunitária desenvolvem-se mediante a acolhida e o compartilhamento comunitário de pequenas inovações que ocorrem em instâncias individuais na interação falante-ouvinte, em grande parte por processos de analogização e *parsing*/neanálise.

<sup>4</sup> Nossa prática de pesquisa centrada em dados da experiência leva-nos a conceber que mudança envolve variação (entendida, em sentido lato, como por similaridade ou por dissimilaridade). Sabemos, entretanto, que algumas práticas de pesquisa construcionistas se voltam para a delimitação de diferenças sincrônica e/ou diacrônica.

<sup>5</sup> A influência de fatores cognitivos assume aqui uma configuração mais ampla que a de Labov (2010), uma vez que entendemos que eles podem afetar não só o que é dito, mas também como é conceptualizado e dito.

e formalmente quanto sobre o processamento, a percepção, a avaliação subjetiva de usos e a atitude em relação a estes. Cabe às generalizações descritivas contemplar o que se passa na Gramática de Construções de uma língua em termos dos fenômenos de estabilidade, variação e mudança. Assumido esse ponto de vista, a questão a ser enfrentada passa a ser a de desenvolver teórico-metodologicamente essa descrição, além da de lidar com o desafio de encarar a questão da variação por similaridade/(quase) sinonímia no modelo da Gramática de Construções centrada na experiência de uso. E, aqui, ainda mais especificamente, procuramos responder à questão de como delinear o envelope da *variação-(meta)construção* nessa proposta de compatibilização teórico-metodológica.

## SOCIOLINGUÍSTICA VARIACIONISTA E GRAMÁTICA DE CONSTRUÇÕES: PERSPECTIVAS DE COMPATIBILIZAÇÃO

### Alinhamento funcional de constructos teóricos

A rede construcional de uma língua conta com padrões construcionais/construções (lexicais ou procedurais/gramaticais) com diferentes graus de esquematicidade (propriedade de generalização de categorias em (sub)esquemas que captam as propriedades formais e funcionais compartilhadas por instâncias), produtividade (frequência *type* e potencial de extensibilidade de um tipo construcional em subtipos) e composicionalidade (relação de transparência entre significação/analísabilidade das partes e significação/analísabilidade do todo construcional, na qual há interferência da contextualidade). Tais padrões mantêm entre si relações de herança e, então, extensão de propriedades. Eles licenciam os constructos/usos (*tokens*) em enunciados orais ou escritos. Podem ser mais esquemáticos, e, então, conter mais *slots* disponíveis para cooptação de unidades da língua (sendo, assim, menos específicos na indicação dos tipos de membros da categoria). Podem ser menos esquemáticos porque são mais preenchidos/substantivos e contam com *slots* mais específicos. E ainda podem ter esquematicidade intermediária. São denominados, respectivamente, de esquemas/macroconstruções, microconstruções, subesquemas/mesoconstruções (cf. TRAUGOTT; TROUSDALE, 2013).

A extensão de uma construção é muito variada: desde unidades lexicais (desde, inclusive, unidades menores que os lexemas) até unidades textuais/textos (cf. HOFFMANN; BERGS, 2018). Ademais, envolve informação sobre unidades linguísticas nela passíveis de compatibilização e regras que presidem a sua

organização (conforme previsto no Princípio de Coerência Semântica, GOLDBERG, 1995). E os *slots* construcionais, a depender do nível de esquematicidade da construção em que se encontram, estarão sujeitos a mais ou menos possibilidades de preenchimento e à compatibilização de membros mais ou menos centrais/prototípicos da categoria.

Logo, as construções sujeitas ao processo de variação as quais aqui focalizaremos podem ter uma destas configurações: (i) ou são unidades/construções da ordem da metaconstrução (conceito a ser definido, a seguir), ou seja, aloconstruções; (ii) ou são unidades/construções da ordem do *slot* construcional, ou seja, (co)lexemas (conceito a ser definido, a seguir). Outra possibilidade de recorte analítico é relacionar ambas configurações, procedendo-se ao cálculo da força de atração/repulsa de um (co)lexema quanto a um *slot* construcional (numa construção, em aloconstruções que configurem metaconstrução), bem como do perfil de construções ou de aloconstruções com que aquele mantém relação de compatibilização (por força de atração/repulsa ou coerção).

Quando se fala em variação no âmbito da Gramática de Construções, a muitos imediatamente ocorre o princípio de não-sinonímia (GOLDBERG, 1995), que prevê a potencialidade, na língua, ou de similaridade semântica ou de similaridade pragmática entre construções diferentes. É, em geral, concebido estritamente com base na relação sintaxe-semântica/pragmática de construções independentes. A remodelagem dessa relação com base nos atributos efetivamente previstos nas faces forma (fonético-fonológica – segmental ou suprasegmental –, morfológica, sintática, lexical) e função/significação (semântica, discursiva, pragmática, social e cognitiva) de um pareamento dá ainda mais margem ao potencial de que, em existindo duas construções ou duas unidades lexicais/lexemas (também construções na teoria) diferentes, haja diferença funcional entre elas. Porém, a feição de dissimilitude não se espalha, necessariamente, a todos os atributos envolvidos nessa face do pareamento (e a diferença, então, é, e não raras vezes, ou apenas semântica, ou discursiva ou pragmática ou social ou cognitiva). Em outras palavras, construções/unidades lexicais formalmente diferentes podem envolver diferença relativa a um atributo ou mais de ordem funcional e não necessariamente a todos os atributos da face funcional. Tal fato enseja, no que diz respeito a alguns atributos, espaço para o mapeamento de comparabilidade funcional, de percepção de similaridade e/ou de associação de construções/unidades lexicais distintas pautada em relativa proximidade funcional. Em decorrência desse processo, duas ou mais construções ((sub)esquemas procedurais ou lexicais) podem ser associadas por força de analogia e, então,

podem ser socialmente rotinizadas e cognitivamente estocadas como alternativas construcionais (aloconstruções); e, ainda, duas ou mais unidades lexicais/construções lexicais têm o potencial de desenvolver o perfil de alternativas no processo de compatibilização a *slot* de uma construção por força de atração (colexemas) ou coerção (lexemas).

Robert (2008), ao lidar, em *Linguística Cognitiva, com princípios de variação e estabilização de redes de unidades lexicais/construções lexicais*, defende esse tipo de abordagem multidimensional<sup>6</sup> na organização entre forma e significação. Variação, inerente a qualquer sistema linguístico, é motivada por mecanismos cognitivos universais (como a analogia, por exemplo) responsáveis pela dinamicidade e plasticidade da conceptualização (por *meaning construal*), uma vez que a significação é construída de várias maneiras (em termos de acesso referencial e de recorte e categorização de conhecimento de mundo), entre (inter) locutores/comunidades diferentes (em termos de relações de poder e de redes sociais, por exemplo) e em contextos muito diversificados (semântica, discursiva, pragmática, social, histórica e culturalmente), mas segundo mecanismos cognitivos comuns. “Intra-linguistic plasticity of meaning echoes inter-linguistic variation” (ROBERT, 2008, p. 55).

---

<sup>6</sup> Concebe uma dimensão na língua para lidar com a complexidade envolvida na relação entre forma e significação/função numa unidade linguística, que ela intitula como: “the depth dimension of language” (“l'épaisseur du langage”, Robert 1999, 2003). “Depth constitutes a third dimension in language, as opposed to the syntagmatic dimension (relations between the utterance’s terms) and the paradigmatic dimension (relations between the terms that may potentially occupy the same spot in the utterance); it is what makes the meaning “subjective and open-ended” (Lichtenberk 1991). This depth dimension constitutes the semantics of a term, and in a way represents the extremely variable harmonics that the semantic-structure-as-fundamental-frequency gives rise to. The depth of language is a complex area where linguistics associates both with linguistic and extra-linguistic matter and which plays an important role in the construal of an utterance’s meaning (Robert 1999).” (p. 73-74) E, na tentativa de ilustrar componentes envolvidos nessa multifuncionalidade, cita processos, que são resumidos pelo editor da obra de que seu artigo (cf. VANHOVE, 2008, p. viii) é parte, como: “This includes cultural categorization, referential paths, internal architecture (figure and ground), metaphor and metonymy, referential scales, application domains, scenarios and semantic universes, networks of formal and semantic relations between terms, connotations and social roles, associations between linguistic and extralinguistic representations, structured relations within a specific verbal or situational context, pragmatic inferences, (de)motivations, landmarks, attractors and “active zones”, prototypes, “semantic isotopics”, and anchor points”.

For both structural and cognitive reasons, natural languages are characterized by their **plasticity**, by **the ease with which the representations borne by the units composing them are subject to change**. Polysemy and polyreference are the general rule among languages. A single unit can thus have several different meanings and point to several different referents. In English for example the word greens can refer to village commons, leafy vegetables or members of a political party. Inversely, different units can refer to the same thing, such as roe and caviar, or hepatitis and jaundice. One could even state that local synonymy (limited to a certain context) is what makes it possible to paraphrase a term or phrase using another. Thus reflect can be paraphrased by either “think” or “throw back light”. **The ability to build equivalences is in fact a fundamental property of language: equivalences between terms (synonymy) or between phrases (paraphrasing), but also between languages (translation)**. There is no one-to-one relation between form and meaning, either within a language or across languages. From this view point, variation within languages (polysemy, synonymy), echoes variation from language to language and raises the question of how it is possible to say “the same thing” differently<sup>7</sup> (ROBERT, 2008: 55-56; grifos nossos).

Construções (unidades procedurais ou lexicais), virtuais por natureza, licenciam constructos/ usos. Porém, também se submetem, na atividade linguística, a operações por meio das quais – além de se atualizarem em enunciados efetivamente produzidos (os quais exibem, em maior ou menor grau, propriedades prototipicamente relacionadas aos padrões construcionais que atualizam) – se compatibilizam noutras construções e/ou se sujeitam aos fenômenos de variação e/ou mudança regular (como as que podem decorrer de situações de ambiguidade ou *misunderstandings* e *mismatches* as quais desencadeiem neoanálise). E, sobre os fatores que promovem variação, atuam, por outro lado, processos de otimização,

---

<sup>7</sup> “Por razões estruturais e cognitivas, as línguas naturais são caracterizadas por sua **plasticidade**, pela **facilidade com que as representações carregadas pelas unidades que as compõem estão sujeitas a mudanças**. Polissemia e polirreferência são a regra geral entre os idiomas. Assim, uma única unidade pode ter vários significados diferentes e apontar para vários referentes diferentes. Em inglês, por exemplo, a palavra *greens* pode referir-se a bens comuns da cidade, vegetais folhosos ou membros de um partido político. Inversamente, unidades diferentes podem referir-se à mesma coisa, como ovas e caviar, ou hepatite e icterícia. Pode-se até afirmar que a sinonímia local (limitada a um determinado contexto) é o que torna possível parafrasear um termo ou frase usando outro. Assim, *refletir* pode ser parafraseado por “pensar” ou “espelhar”. A capacidade de construir equivalências é de fato uma propriedade fundamental da linguagem: equivalências entre termos (sinonímia) ou entre frases (paráfrase), mas também entre idiomas (tradução). Não existe uma relação biunívoca entre forma e significado, dentro de um idioma ou entre idiomas. Desse ponto de vista, a variação nas línguas (polissemia, sinonímia) ecoa a variação de idioma para idioma e levanta a questão de como é possível dizer “a mesma coisa” de maneira diferente.”

regulação e estabilização de significação. Isso significa que “language is the seat of tensions between opposing forces which can all be functionally justified”<sup>8</sup> (ROBERT, 2008, p. 58).

É preciso considerar, por fim, que a significação/referência linguística é mediada por “usos linguísticos”, uma vez que unidades lexicais e/ou enunciados usados não são as entidades/os estados de coisas (dinâmicos ou não) do mundo biossocial ou imaginário. Expressões linguísticas materializadas na comunicação constituem mecanismos de apreender representações de uma realidade (qualquer que seja a configuração desta) que é sentida, percebida e/ou concebida com base numa rede complexa de relações, bem como elementos “vivos”, oriundos das capacidades linguísticas do falante de repetir e criar, no jogo complexo de forças de estabilidade e instabilidade, um jogo discursivo-pragmático, social e histórico, culturalmente situado.

Languages therefore show equivalency relations, although construals and reference constructions are extremely variable. (...) **To gain access to a same referent, languages construct variable reference pathways. (...) The variability of referential paths across languages, as well as inside a given language,** is due to a more general property of language, as claimed by cognitive linguistics, namely its **ability to “construe” a particular situation in different ways** (LANGACKER, 1991a<sup>9</sup> *apud* ROBERT, 2008, p. 59).

Usos observados em pesquisa linguística revelam-se, então, medidas (discretas, segmentadas pelo pesquisador numa realidade complexa) de padrões construcionais (lexicais ou procedurais) que se podem apresentar estavelmente associados por similaridade/analogia num recorte semântico-conceitual, discursivo-pragmático, social, temporal, cultural.

---

<sup>8</sup> “a linguagem é a sede de tensões entre forças opostas que podem ser justificadas funcionalmente”.

<sup>9</sup> “As línguas, portanto, mostram relações de equivalência, embora *construals* e construções de referência sejam extremamente variáveis. (...) **Para obter acesso a um mesmo referente, as línguas constroem caminhos de referência variáveis. (...) A variabilidade dos caminhos referenciais entre as línguas, bem como dentro de uma determinada língua,** é devida a uma propriedade mais geral da linguagem, conforme reivindicada pela linguística cognitiva, ou seja, sua **capacidade de “interpretar/conceber” uma situação específica de diferentes maneiras.** (LANGACKER, 1991a)

## O DESENHO DO ENVELOPE DA VARIAÇÃO DE PADRÕES CONSTRUCIONAIS NUMA METACONSTRUÇÃO

As variantes de uma área de sobreposição semântica/funcional, em que se prevê uma alternância/variação, são denominadas, segundo Cappelle (2006), como *allostructions* (aloconstruções), já que elas correspondem a possibilidades de representação configuracional/esquemática de uma (ou mais) propriedade(s)/atributo(s) de funcionalidade/significação. Por exemplo, o evento de transferência de posse, no inglês, pode ser viabilizado por duas estruturas (*the dative alternation*): *SN V SN SN (the ditransitive pattern)* ou *SN V SN SP (the oblique-goal pattern)*. O autor define *allostructions*, por analogia a *alofone* e *alomorfe*, como possibilidades alternativas – padrões construcionais em relação de variação – de uma unidade linguística parcialmente especificada: “variant structural realizations of a construction that is left partially underspecified”<sup>10</sup> (CAPPELLE, 2006, p. 18).

Perek (2015) argumenta que se pode representar, na rede construcional, generalização de um significado comum (ou de uma zona de funcionalidade semântica, discursiva, pragmática, social ou cognitiva partilhada) entre dois ou mais padrões construcionais independentes. Tais padrões são, assim, associados a uma forma não específica (subespecificada), em razão de similaridades observáveis. Para sustentar a relação de associação entre construções, Cappelle (2006) argumenta que “duas” construções devem ser concebidas como ligadas mediante uma “metaconstrução”, constructo teórico-descritivo que capta o que essas construções apresentam em comum e uma (relativa) neutralização do que as diferencia/faz unidades diferentes.

Perek (2015) sugere a denominação *constructeme* (correspondente a metaconstrução)<sup>11</sup>. A representação por metaconstrução e dos seus respectivos “links” de herança capta, de um lado, a semelhança entre padrões construcionais e abarca, de outro lado, nas *allostructions*, as especificações semântico-pragmáticas (e cognitivas e sociais) adicionais, bem como o tipo de informação funcional prototípico de cada aloconstrução. Em outras palavras, *metaconstrução* captura o nível de representação em que construções são funcionalmente equivalentes; e *aloconstrução* explicita exatamente como tais construções diferem (por quais

<sup>10</sup> “Realizações estruturais de uma construção que é deixada parcialmente subespecificada” (CAPPELLE, 2006, p. 18).

<sup>11</sup> Perek (2015, p. 153) menciona que Cappelle também chama a generalização que envolve aloconstruções de *constructeme*.

valores de atributos). Assim, a modelagem por *metaconstrução com aloconstruções* apreende a relação entre similaridade e dissimilaridade (ou seja, de valores, nas construções, ligados a atributo(s) destas – semântico(s), pragmático(s), discursivo(s), social(is) ou cognitivo(s)).

Para exemplificar o alinhamento dos conceitos de aloconstrução e metaconstrução ao dos conceitos envolvidos na configuração do envelope da variação, valemo-nos dos resultados da pesquisa realizada por Silva, Fontenlos e Justen (2017), sintetizada a seguir.

Ao investigar a variação na flexão (singular ou plural) dos verbos modais *poder* e *dever* em locuções verbais constituídas de verbos principais transitivos diretos, as autoras examinaram dados do uso que instanciassem o padrão construcional [V<sub>auxiliar</sub> V<sub>Transitivo Direto</sub> + SE<sub>Apassivador/Indeterminador de Participante</sub> SN]. Então, categorizaram constructos/ usos reunidos na amostra que estivessem vinculados a estes dois padrões construcionais: (1) [V<sub>aux singular/plural</sub> V<sub>TD</sub> + SE<sub>Apass./Indet. de Part.</sub> SN<sub>Não-ativo singular/plural</sub>] e (2) [V<sub>aux singular</sub> V<sub>TD</sub> + SE<sub>Indet. de Part.</sub> SN<sub>Não-ativo singular ou plural</sub>]. Com o primeiro, tencionaram captar a possibilidade de variação na flexão em sintonia com SNsingular ou SNplural; com o segundo, procuraram representar a estabilidade da flexão verbal singular, estando o SN quer no singular quer no plural<sup>12</sup>. São os constructos licenciados por esses padrões construcionais que herdam propriedades da construção mais esquemática [V<sub>auxiliar</sub> V<sub>Transitivo Direto</sub> + SE<sub>Apass./Indet. de Participante</sub> SN<sub>não-ativo</sub>]:

Ex. 1 “Com isso, ***podem-se analisar*** também as tendências do jornalismo científico capixaba no que diz respeito à valorização de setores da ciência e do uso da linguagem para se falar de pesquisa.” (www.jornalismocientifico.com.br/revista/01/artigos/artigo1.asp) (1)

Ex. 2 “Diferente do modo convencional, ***pode-se explorar*** muitas ideias relacionadas a números. Ao brincar as crianças enfrentam desafios e problemas, buscando soluções para resolvê-los.” (https://jornalhoraextra.com.br/.../4658-brincar-e-importante-para-desenvolvimento-das-crianças, 11 de out de 2017) (2)

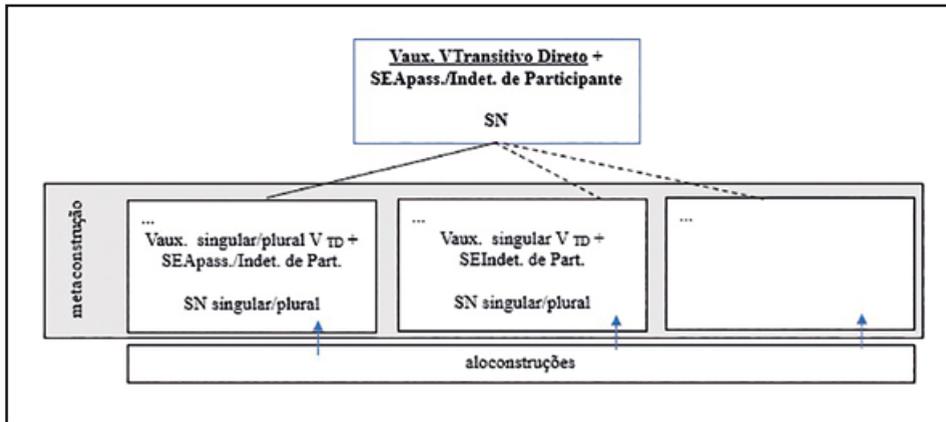
<sup>12</sup> “(...) o argumento único do predicado na construção impessoal se comporta como Sujeito real: além de ocupar uma posição destinada ao Objeto, nem sempre se mantém a codificação morfossintática que regula o comportamento nominativo desses SNs. Cria-se, assim, uma espécie de voz ativa impessoal indeterminadora, em que o argumento Paciente não recebe função de Sujeito, cuja posição fica marcada formalmente pela presença do clítico se” (CAMACHO, 2000, p. 218).

Ex. 3 “No caso de não surtir efeito, pode-se acionar a Justiça para pleitear a indenização por danos morais.” (<http://www.gazetadopovo.com.br/justica/cobrar-dividas-de-forma-abusiva-e-ilegal-e-pode-ate-dar-cadeia-cp53kyqsud8x5r8oag2qblfg1>) (1) ou (2)

No exemplo 1, o verbo poder está no plural acompanhando o número do SN (“as tendências do jornalismo científico capixaba”), o que nos leva a cogitar de uma relação entre esse enunciado e o padrão construcional (1) – referido em parágrafo imediatamente antes dos exemplos. Já no exemplo 2, está no singular, não acompanhando o número do SN “muitas ideias relacionadas a números”, o que nos faz supor uma relação entre esse enunciado e o padrão construcional em que o verbo fica sempre no singular (padrão construcional 2). Por fim, no exemplo 3, está no singular, acompanhando o número do SN “a Justiça” ou não, o que nos faz suspeitar da possibilidade de o enunciado se caracterizar como licenciado pelo padrão construcional (1) ou pelo padrão construcional (2).

Entre suas generalizações sobre estratégias para assinalar opacificação/(des) focalização de um participante por impessoalização discursiva (e, assim, não perfilar um participante – no caso, indutor/trajetor do estado de coisas, agente/causador/força/experienciador –, tornando-o implícito e voltando o foco para o evento/estado em si e/ou outro participante – o SN não-agentivo/não-indutor, meta/marco – na proposição), o Português possui a variação por similaridade configuracional entre certas expressões que instanciam padrões construcionais de uma construção mais esquemática de predicação verbal pessoal com pronomes SE passivador/indeterminador de participante. No discurso científico, podemos encontrar constructos de dois padrões gramaticais, um com concordância (verbo e SN<sub>Não-agentivo</sub> sujeito) e outro com verbo sem concordância (verbo sempre no singular). Representamos essa variação por semelhança configuracional (em que os padrões revelam partes estáveis e atributos similares, até por causa da relação de herança por subparte a que se submetem, e atributos diferentes) no esquema 1, a seguir:

Esquema 1: Representação de aloconstruções e metaconstrução<sup>13</sup>



No esquema (1), representamos a construção  $[V_{\text{aux.}} V_{\text{Transitivo Direto}} + SE_{\text{Apassivador/Ideterminação de Participante}} SN]$ , na parte superior, nível mais abstrato, isto é, mais esquemático, que gera, no Português Brasileiro, possibilidades (alternâncias) de realizações linguísticas:

- i) com item verbal auxiliar com flexão moldada pelo número do SN a que se conecta sintaticamente a locução verbal ou (mais ou menos) indiferente a este;
- ii) em configuração de estruturação passiva, com indeterminação de participante e com  $SN_{\text{Não-agentivo}}$  sujeito, ou de estruturação ativa, com indeterminação de participante e com  $SN_{\text{Não-agentivo}}$  complemento.

Elas são licenciadas por aloconstruções, que, no esquema 1, se apresentam como representações, ainda esquemáticas, de “possibilidades de realizações” com alta similaridade configuracional. Ambas aloconstruções partilham propriedades da construção mais esquemática ( $[V_{\text{aux.}} V_{\text{Transitivo Direto}} + SE_{\text{Apassivador/Ideterminação de Participante}} SN]$ ), pois possibilitam leitura de indeterminação de participante envolvido no evento conceptualizado, sendo variantes/alternativas de expressão em alguns contextos de predicação, embora em outros possam não ser. Assim

<sup>13</sup> O último retângulo no espaço em cinza, espaço de representação da metaconstrução, deixado apenas com reticências sinaliza que a construção mais esquemática, representada no topo do esquema, licencia outros padrões construcionais, que apenas não interessam à exemplificação em foco neste momento. A linha contínua indica uma ligação com um exemplar construcional (uma das aloconstruções) caracterizado como prototípico de  $[V_{\text{aux.}} V_{\text{Transitivo Direto}} + SE_{\text{Apassivador/Ideterminação de Participante}} SN]$ . A linha pontilhada indica que um padrão construcional viável (a outra aloconstrução) não é caracterizado prototipicamente como tal.

sendo, o conceito de variação liga-se à área de intersecção de ambas aloconstruções representadas pela área em cinza, a metaconstrução.

Pensando em alguns conceitos clássicos da Sociolinguística Variacionista (variante e variável), podemos alinhar: (a) aloconstruções como *variantes* e (b) metaconstrução, grosso modo, como significado referencial, função semântico-discursivo-pragmática, função/significação social, função cognitiva (de conceptualização)<sup>14</sup> partilhado(a) ou *macrofunção* (uma espécie de “*arquifunção*”<sup>15</sup>) e, enfim, como *variável*<sup>16</sup>.

Assim, as variantes construcionais menos esquemáticas de impessoalidade discursiva (por indeterminação de participante) partilham uma parte significativa de seu significado/de sua configuração forma-função e, conseqüentemente, são estocadas no conhecimento linguístico como duas alternativas construcionais para a expressão de estados de coisas com a opacificação/(des)focalização de sua força indutora<sup>17</sup> (agente, causador, força, experienciador; que, inclusive, pode coincidir com o enunciador). As variantes construcionais têm também, entre seus atributos, propriedades formais e funcionais apreendidas em tendências (estatísticas) correlacionadas a fatores-valores de atributos que governam a instanciação no uso de uma ou de outra: modalidade expressiva, grau de generalização da predicação, distância do SN da locução verbal e ordem daquele em relação a esta.

Dessa forma, as aloconstruções (*allostructions*) designam, então, os padrões construcionais variantes em si enquanto a metaconstrução é a nomenclatura usada para captar o espaço/domínio de generalização partilhado e configurado em uma área de representação abstrata da neutralização/opacificação das diferenças desses padrões. Esses construtos teóricos serão importantes para lidar com fenômenos variáveis e para a caracterização do envelope da variação como o aqui brevemente descrito, pois (i) o primeiro possibilita compreender generalizações sobre padrões construcionais que vão além das relações de instanciação e (ii) o segundo, metaconstrução, captura o que resulta de relações analógicas e do uso criativo na linguagem. Conforme Leino e Östman (2005) salientam, metaconstrução não pode confundir-se como um “nível” de padrão construcional

<sup>14</sup> A depender da perspectiva de análise e descrição do objeto variável em foco (por exemplo, Sociolinguística; Sociolinguística e Funcionalismo; Sociolinguística e Cognitivismo).

<sup>15</sup> Na linha de *arquifonema*, *arquimorfema*.

<sup>16</sup> Variável dependente, na rodada multivariada (normalmente feita via Goldvarb X).

<sup>17</sup> Indutor é uma nomenclatura acionada aqui para designar em sentido lato um dispositivo ativo ou passivo (neste segundo caso, como no mundo da elétrica).

mais abstrato, geral ou esquemático (“subesquema construcional” ou “meso-construção”), pois abrange não só semelhanças sistemáticas, mas também diferenças que ocorrem entre padrões construcionais; só que, ademais, representa o resultado de um processo de “neutralização/opacificação” dessas diferenças em razão de pensamento analógico. E esse resultado é uma generalização também estocada na mente do falante<sup>18</sup>.

Se lembrarmos que o usuário da língua produz enunciados/instâncias licenciados pelas construções no seu dia a dia – justamente o espaço em que há, em potencial, *matches* ou *mismatches* e *misundertandings*, bem como extensões de uso – e considerarmos que o faz com base em restrições e tendências estatísticas (ocorre-nos aqui a noção de “preempção estatística”, BOYD; GOLDBERG, 2011, e GOLDBERG; BOYD, 2015<sup>19</sup>), passamos, então, a configurar a noção de aloconstrução considerando ainda os parâmetros de produtividade e esquematicidade. Os padrões construcionais *variantes* que figuram numa metaconstrução podem ser condicionados por *variáveis independentes* (ou seja, “condicionamentos internos ou externos” na visão sociolinguística) aqui denominada(o)s de *atributos*, vinculadas às faces forma e função de uma construção. E isso nos remete ao “problema das restrições” (WEINREICH; LABOV; HERZOG, 1968). Com base no *know-how* já alcançado no âmbito da Sociolinguística Variacionista (em suas diferentes práticas teórico-metodológicas e analíticas), concebemos restrições sob duas perspectivas: geralmente, com base em *valores/fatores* de *atributos/variáveis independentes* que se materializam em tendências estatísticas de uso variáveis, porque abaixo de 100%, ou com base em *valores/fatores* de *atributos/variáveis independentes* que se materializam de forma categórica.

Tais noções (*atributo/variável independente, valor/fator, condicionamento variável/não-categórico ou condicionamento categórico*) passam a orientar

---

<sup>18</sup> Além de generalizações sobre a existência independente das construções envolvidas nessa relação de similaridade.

<sup>19</sup> GOLDBERG; BOYD (2015, p. 185): “**Statistical preemption** or “blocking” of a target form is the process of learning to avoid a potential target form because a competing form has been consistently witnessed instead in contexts in which the target form would otherwise have seemed appropriate. This is the widely accepted view of how children learn to avoid morphological overgeneralizations such as *goed: went* is consistently witnessed in contexts in which *goed* might have been expected to occur (Aronoff 1976; Kiparsky 1983). A number of researchers have proposed that statistical preemption is capable of scaling up to account for certain non-occurring syntactic formulations that have readily available competing alternatives (Ambridge et al. 2012, Brooks & Tomasello 1999, Goldberg 1995, Goldberg 2006, Goldberg 2011c, Payne et al. 2013, Poser 1992, Robenalt & Goldberg to appear)”.

a representação da formulação das expressões linguísticas no uso (variantes) e, então, precisam ser captadas na rede construcional. De acordo com Leino e Östman (2015), podemos representar, nas construções – que envolvem, na sua configuração, atributos –, o que é denominado pelos autores de “*value pool*” (valores de atributos), contando, para tanto, com um tratamento quali-quantitativo de dados, como na Sociolinguística.

Para ilustrarmos a correlação de atributos a fatores-valores na configuração de aloconstrução e metaconstrução, recorreremos aos resultados da investigação conduzida por Alves (2011), que analisou dados de predicação constituídos por predicadores simples (verbos plenos acompanhados de clítico SE) e predicadores complexos (construção com verbo suporte), conforme os exemplos:

*Ex. 4 “E quando ele viu o que Gabriel estava fazendo, **se desesperou**, dizendo.” [9º E.F, narração] – verbo pronominal*

*Ex. 5 “Luíza **ficou desesperada** e cada vez mais estava perto do fundo do poço, teve seu filho, e acabou na rua, pedindo esmola.” [9º E.F, narração] – perífrase com verbo suporte*

O processo de comparabilidade funcional entre verbo pronominal exemplificado no primeiro enunciado (*se desesperou, preocupar-se*), e predicadores complexos com verbo suporte e elemento não-verbal cognato a lexema verbal simples/pleno (*ficou desesperada, ter preocupação*) foi estudado por Alves (2011). A pesquisadora partiu de uma afirmação de Kato (1998, p. 233) de que, em decorrência de mudanças no quadro pronominal, os falantes tenderiam a um maior emprego de verbos leves (*ficar interessado*) como forma de evitarem o clítico SE dos verbos reflexivos inerentes (*interessar-se*). E também partiu da hipótese, com base em pesquisas de experiências de uso e de percepção/avaliação subjetiva de uso do projeto PREDICAR (entre as quais, ESTEVES, 2008), de que tal processo revelaria uma situação estável de alternância (variação construcional, no enfoque explorado neste capítulo) entre possibilidades estruturais relacionadas por similaridade de funcionamento na predicação de um estado de coisas, situação em que, segundo a autora, poderiam interferir (a) atributos/grupo de fatores formais (posição do predicador no período, estatuto do pronome SE, extensão silábica do predicador, posição do SN<sub>Não-agentivo</sub>, por exemplo) e (b) atributos funcionais (relativos ao tipo de organização discursiva, à modalidade discursiva, à intenção do enunciador, ao ponto focal/proeminente na conceptualização de um estado de coisas, ao nível de escolaridade dos informantes, à caracterização semântica do estado de coisas e do participante SN<sub>Não-agentivo</sub>, por exemplo).

Alves (2011) observou, entre as estruturas com verbos acompanhados de pronome SE, estruturas pronominais em que uma variante possível/virtual é a que envolve verbo suporte (geralmente, *ficar*, conforme exemplo 5, ou *ter*<sup>20</sup>), como materializadas, respectivamente, nestes exemplos da amostra por ela examinada:

Ex. 6 “Quanto a cota racial não vale a pena nem discutir, sou totalmente contra, porque nem todo pobre é negro e nem todo negro é pobre, sem contar que o preconceito iria continuar existindo, pois imagina uma sala com 5% de negros, 2% de índios e todo resto de branco? Sem nenhuma dúvida os índios e os negros não **iriam se adaptar** e de nada valeria o sistema de cotas.” [3º ano E.M, dissertação] ~ {**teriam adaptação**}

Ex. 7 “Mas não podem afirmar como verdade absoluta, afinal não há somente uma sociedade com as mesmas regras e costumes onde há ou não oportunidades para a mulher **se destacar**.” [3º ano E.M, dissertação] ~ {**ter destaque, ser destacada**}

Segundo a autora, “o uso da forma simples não causa o mesmo efeito: a ênfase na ação praticada por outros se perde, pois a forma simples dá a impressão de que os próprios participantes/sujeitos que sofrem a ação são responsáveis por ela”. Isso pode ser percebido respectivamente nos excertos abaixo:

Ex. 8 “Os R\$ 0,20 foram a gota d’água. O copo e a nossa paciência transbordaram. As reivindicações **se nacionalizaram, se internacionalizaram**, tomaram conta das ruas, das redes sociais, dos jornais (mesmo a contragosto de alguns), do congresso, das conversas... (Protestos: a gota d’água, <http://democraciapolitica.blogspot.com/2013/06/protestos-gota-dagua.html>)

Ex.9 “Uma carreira acadêmica pode ser destruída por universidades ou governos com base no fato de que o conteúdo de um trabalho, seja real ou imaginado, é uma possível ameaça aos poderes existentes. Pode ter sido um programa de um curso ou o tema de uma dissertação supervisionada que tenha despertado a ira do Estado; ou talvez as posições políticas tomadas dentro ou fora dos muros da universidade – sindicalização, desmilitarização, oposição ao nacionalismo. Tais posições são distorcidas pelos censores e por aqueles com poder para destruir uma carreira e exilar um cidadão. As reais posições da pessoa são exageradas, demonizadas e objeto de sensacionalismo. Um apelo à democracia é interpretado como insubordinação; um pedido de paz **transforma-se** em uma aliança com o terrorismo; um pedido de liberdade é considerado um chamado à violência.” (A criminalização do conhecimento, <http://sxpolitics.org/ptbr/a-criminalizacao-do-conhecimento-por-judith-butler/8391>)

---

<sup>20</sup> “Neymar **teria adaptação** mais fácil no Real Madrid, analisa Ledio Carmona.” (<https://www.sidrolandianews.com.br/noticia/brasil/neymar-teria-adaptacao-mais-facil-no-real-madrid-analisa-ledio-carmona>).

Uma particularidade a destacar quanto ao alinhamento entre a noção de fatores/variáveis independentes e a de valores/atributos consiste no fato de que a noção deste constructo teórico, diferentemente da noção de condicionamento/restrrição, engloba, ainda, a potencialidade de ele atuar com força de coerção<sup>21</sup>, e não só com forças de atração/repulsa e restrição<sup>22</sup>. Cada uma dessas possibilidades de forças atuantes terá de ser considerada na investigação socioconstrucionista, a depender de como se configure o fenômeno variável em pauta<sup>23</sup>.

Em linhas gerais, o mapeamento dos padrões construcionais/aloconstruções que viabilizam opacificação/(des)focalização de participante e/ou impessoalização discursiva leva-nos à identificação de padrões construcionais que, por um lado, podem ser associados, por similaridade, em termos do estado de coisas que conceptualizam e, por outro lado, diferem em termos de como predicador e participantes são perspectivados e perfilados na configuração linguística do estado de coisas<sup>24</sup>. Assim, pesquisas (como as citadas) consideram<sup>25</sup> a potencialidade

<sup>21</sup> “Coercion is achieved by imposing the application of a construction even though not all its criteria are satisfied” (STEELS, 2013, p. 167). Coerção é um tipo de força de acomodação ou reinterpretção contextual impulsionada pela necessidade de resolver/reconciliar uma combinação semântica parcial, discrepante ou em conflito (um *mismatch*) entre uma unidade construcional e outra mais esquemática.

<sup>22</sup> Ao leitor interessado sobre o assunto, indicamos a leitura de Wiedemer e Machado Vieira (2018b).

<sup>23</sup> E, com base nesse aspecto, já se antecipa uma diferença entre Sociofuncionalismo e Socioconstrucionismo (cuja comparação procuramos esboçar adiante neste capítulo). De certo modo, podemos dizer que aquela perspectiva põe em evidência as forças de atração/repulsa e restrição na análise da coatução de variáveis sobre usos variáveis ou (semi)categoricos (LABOV, 2003; segundo o qual haveria regras linguísticas *variáveis/com* frequência na faixa de 5% a 95%, *semicategoricas/com* 95%-99% de frequência e *categoricas/com* 100% de aplicação), e não as forças de coerção de uma representação (sub)esquemática e/ou de compatibilização, também atuantes na língua segundo a concepção construcionista.

<sup>24</sup> Há sempre uma tensão de forças que coatuam quando se concebe variação construcional em moldes multidimensionais e multifatoriais. “Ser um conceito multifatorial significa que a voz verbal representa um grande número de valores e de possibilidades correspondentes de expressão que, segundo Givón (1981, 1994), envolvem três domínios funcionais: a) topicalidade: atribui-se a função de Tópico a um argumento não-Agente; esse comportamento é oposto ao da sentença ativa correspondente, em que o Tópico é comumente o Sujeito/Agente; b) impessoalidade: suprime-se a identidade/presença do argumento Agente, geralmente o Sujeito expresso da sentença ativa; c) detransitividade: a construção de voz é semanticamente menos “ativa”, menos transitiva, mais estativa que a construção “ativa” correspondente” (CAMACHO, 2000, p. 216).

<sup>25</sup> Além de Alves (2011), outro estudo sobre passivas (em curso no mesmo projeto de pesquisa – Projeto PREDICAR –, mas agora sob um olhar socioconstrucionista) é o de MACHADO

de variação construcional entre, por exemplo: (i) o padrão de passiva sintética sem perfilamento de indutor/trajetor (agente, causador, força, experienciador), na verdade, com a suspensão deste; (ii) o padrão de passiva analítica sem perfilamento de indutor (agente, causador, força, experienciador) ou com perfilamento de indutor por meio de marcas de indefinição/opacificação de seu referente; (iii) o padrão de estrutura ativa com perfilamento de indutor por meio de mecanismos de indefinição/opacificação de seu referente. Por exemplo:

(i)

Ex. 10 “Foi só a partir da afirmação do poder local democrático que **nas diversas localidades disseminadas pelo território nacional se rasgaram caminhos, pavimentaram-se quilómetros de vias, quebrou-se o isolamento de populações, eletrificaram-se aldeias, distribui-se a água, implantou-se o saneamento, edificaram-se pontes, construíram-se escolas, colocaram-se a funcionar bibliotecas, museus, pavilhões, piscinas, multiusos, centros de dia, multiplicaram-se associações, enfim sobreveio um pulsar vigoroso da sociedade portuguesa. (O pelouro da diáspora nas autarquias, <http://www.palopnews.com/index.php/home/99-noticias-gerais/2415-o-pelouro-da-diaspora-nas-autarquias>)**

(ii) a

Ex. 11 “Como sabemos, os verdadeiros pontos de vista políticos pelos quais **os académicos são punidos** são aqueles direcionados para as políticas de um governo ou para uma universidade que tenha práticas injustas, formas de exploração, que use serviços de segurança e vigilância para acabar com questionamentos abertos e discussões públicas, ou uma universidade que tenha vínculos com interesses de Estado ou interesses corporativos resultando em um controle do corpo docente.”. (Sinalização de trânsito é retirada na DF 079, <http://sxpolitics.org/ptbr/a-criminalizacao-do-conhecimento-por-judith-butler/8391>)

(ii) b

Ex. 12 “Agentes de trânsito colocaram sinalizadores para evitar o congestionamento, mas **foram retirados por alguém**” (Sinalização de trânsito é retirada na DF 079, <http://www.jornaldebrasil.com.br/conteudo-tv-record/sinalizacao-de-transito-e-retirada-na-df-079/>)

(iii)

Ex. 13 Quando os ventos da mudança sopram **alguns constroem muros. Outros constroem moinhos de vento.** O goleiro Ricardo Prasel parou de jogar precocemente sofrendo com dores crônicas no quadril e achou que nunca mais alcançaria o estrelato através do espor-

---

VIEIRA, SANTOS; KROPF (2019). Vale ressaltar, ainda, que Alves (2011) atesta a viabilidade da comparabilidade funcional entre passiva sintética e passiva analítica com base em testes de avaliação subjetiva.

te. (<https://noticias.bol.uol.com.br/ultimas-noticias/esporte/2017/08/25/ex-goleiro-do-chelsea-paranaense-vira-peso-pesado-no-mma-e-sonha-com-ufc.htm>)

Na projeção que fazemos quanto à configuração dessa rede complexa de padrões construcionais de predicação para pôr em evidência passividade, cogitamos, inclusive, de outras possibilidades de pareamento, como o que licencia estruturas semelhantes à que destacamos no excerto abaixo (com “sofrer”, [sofrer SN]<sub>predicador complexo</sub>):

*Ex. 14 “Vamos ponderar a diferença entre liberdade acadêmica e direito de expressão política, que, como Joan Scott deixa claro, não são a mesma coisa. A liberdade acadêmica pertence ao corpo docente de uma universidade que foi nomeado para ensinar, buscar e produzir conhecimento. A expressão política é o direito do cidadão de expor seu ponto de vista político da maneira que quiser. As duas coisas convergem quando acadêmicos que falam em espaços “extramuros” sofrem retaliação ou punição dentro da universidade ou são ameaçados de perder seus cargos.” (A criminalização do conhecimento, <http://sxpolitics.org/ptbr/a-criminalizacao-do-conhecimento-por-judith-butler/8391>)*

## O DESENHO DO ENVELOPE DA VARIAÇÃO DE (CO)LEXEMAS EM *SLOT* CONSTRUCIONAL

Outra situação linguística é a que enseja (i) variação de *colexemas* (collexemes<sup>26</sup>) por força de atração de certas unidades linguísticas a um *slot* construcional ou (ii) variação de lexemas por força de coerção da construção sobre as formas que são acionadas para preenchimento do *slot* construcional<sup>27</sup>. No primeiro tipo, as possibilidades de preenchimento atraídas ao *slot* na construção partilham propriedades formais e/ou funcionais. Por isso, são rotuladas de *colexemas*. *Colexemas* são unidades lexicais funcionalmente compatíveis com o significado construcional. No segundo tipo, as unidades assumem uma certa configuração formal e/ou funcional similar ao de outras unidades (mais atraídas, porque compatíveis como as do primeiro tipo) por força de imposições advindas da construção. As

<sup>26</sup> Palavras que são atraídas para uma construção particular são referidas como *colexemas* dessa construção (STEFANOWITSCH & GRIES, 2003, p. 214–215).

<sup>27</sup> A título de ilustração, vale lembrar que o verbo “ser” é normalmente o verbo mais atraído para o *slot* construcional de verbo auxiliar de voz passiva, embora outros verbos (*estar*, por exemplo) possam ser “forçados” a figurar nele com esse papel. Por exemplo, “são obrigados a aguentar a falta de educação desses “princesas” e “príncipes” que *estão criados por* pais que não respeitam professores (...).” (<https://jovempan.uol.com.br/arquivo/marilena-chauai-firma-que-quem-defende-familia-e-uma-besta.html>)

do segundo tipo são, por sua vez, referidas como *lexemas*. Em outras palavras, a um *slot* construcional podem-se compatibilizar: (i) unidades com similaridade formal-funcional, isto é, alternativas atraídas pelo *slot* categorial/*colexemas* e, então, mais esperadas em razão das condições de preenchimento implicadas por ele e, naturalmente, de inclinações de uso numa comunidade; e (ii) unidades menos ou não sintonizadas às condições do *slot* categorial e, então, não ou menos atraídas pelo *slot/lexemas*, que, por força de coerção advinda da construção, são forçadas a se alinharem formal e funcionalmente às mais frequentemente compatibilizadas em tal *slot* e, então, passam a atuar como alternativas a essas. Os (co)lexemas podem revelar-se, então, como variantes de uma *variável dependente/um slot construcional*, que conta com restrições/condicionamentos de preenchimento.

Já é sabido na literatura construcionista que a compatibilização funcional não é o único fator que afeta a relação entre *lexemas* e *slots* construcionais. Outro fator é, sem dúvida, sua frequência de uso, seu entrincheiramento na memória. Entendemos que a experiência de uso numa comunidade também influencia a relação/associação entre unidades lexicais e construção cognitivamente rotinizada e estocada para facultar propósitos sociocomunicativos de conceptualização e expressão no mundo.

Sobre o assunto, Hilpert (2014) sugere que a análise de (co)lexemas e sua relação com a construção nos permite alcançar a percepção de que o significado de uma construção tende a se harmonizar com os significados dos elementos lexicais que normalmente ocorrem nela. Dessa forma, podemos compreender que a variação a partir da atuação das propriedades (de preenchimento) do *slot* construcional, seja por força de coerção, seja por força de atração/repulsa de *lexemas* em razão da produtividade/rotinização do uso na construção, seja por força de restrição.

Por exemplo, retomando a pesquisa de Alves (2011), nem sempre determinados *lexemas* podem ocorrer nas duas variantes construcionais, pois podem existir algumas combinações de verbos instrumentais/gramaticais às quais há restrição; ou, ainda, a ordenação pode restringir a seleção de *lexemas* para os *slots* envolvidos na construção, o que nos permite enxergar a variação em conjunto com a relação entre *lexema* e construção gramatical. Observando, por exemplo, os usos verbais envolvidos na amostra de dados investigada pela autora, percebe-se que o maior número de perífrases ocorre com os itens *estar* (8%, 22/260), *ser* (30%, 78/260) e *ficar* (33%, 87/260), de um total de 260 ocorrências de perífrases analisadas (além de perífrases verbo-nominais envolvendo também outros itens

verbais, como “*ter*”, “*sofrer*”, “*tomar*”, “*levar*”). Além disso, o verbo *ser* aparece como auxiliar de passiva analítica; o verbo *estar* aparenta funcionalidade típica do verbo instrumental em predicados nominais (verbo relacional, conforme descrevem PAVÃO; MACHADO VIEIRA, 2013), que Ranchhod (1990) classifica como verbo-suporte e Dik (1997) rotula de verbo cópula suporte. E *ficar* também se alinha a membros das categorias de verbo suporte e verbo auxiliar de passiva; ademais se compatibiliza normalmente em padrões de construção predicativa relacional que implicam ou mudança de estado ou mudança de propriedade (conforme descrito por FERREIRA, 2019). Os resultados do estudo sociofuncionalista de Alves (2011) oferecem-nos indícios, conforme já apontamos, de que o significado de uma construção tende a se harmonizar com os significados dos elementos que normalmente ocorrem nela (*ser*, *estar* e *ficar* são (co)lexemas em variação no *slot* destinado a verbo auxiliar de voz passiva); daí a necessidade de se investigar a relação entre lexemas e construções.

Assim, em contraste à abordagem discursivo-funcional ou ao Sociofuncionalismo, que têm focado suas pesquisas nas “diferenças” de ordem funcional entre diferentes padrões de usos, no modelo assumido aqui as propriedades formais e funcionais são vistas a partir de concepção de pareamento forma-função/significado, em que a natureza do relacionamento dos diferentes atributos abarcados por esses dois polos se torna importante para a investigação.

Além disso, Gries e Stefanowisch (2004) destacam que, na abordagem da Gramática de Construções, por exemplo, Goldberg (2002) toma a “alternância” como uma espécie de “paráfrase parcial” quando usada com certos itens lexicais e mostra “that a construction containing a given verb shares more semantic and syntactic properties with constructions of the same kind containing different verbs than with the other construction of the alternating pair containing the same verb”<sup>28</sup>. Porém, para os autores,

In the context of alternating pairs, a focus on constructional semantics and semantic compatibility raises several questions: first, what exactly are the (often seemingly tenuous) semantic differences between the members of such a pair; second, how productive is the ‘alternation’ in actual usage, i.e. which verbs/nouns occur freely in both constructions, and which have strong biases towards one of them; and third, is a constructional, non-derivative approach plausible given the answers to the first two questions. We believe that a method that extends the notion of distinctive collocates in the context of our

<sup>28</sup> “que uma construção contendo um verbo compartilha mais propriedades semânticas e sintáticas com construções do mesmo tipo contendo diferentes verbos do que com a outra construção do par alternado contendo o mesmo verbo”.

previously proposed ‘collostructional analysis’ may provide answers to these questions.<sup>29</sup> (GRIES; STEFANOWISCH, 2004, p. 99).

## O DESENHO DO ENVELOPE DA VARIAÇÃO: UM PANORAMA GERAL E POTENCIALIDADES FUTURAS

Voltando ao caso da construção [predicação verbal com desfocalização/opacificação da força indutora do estado de coisas (dinâmicos ou não)], vista na seção anterior, podemos categorizar os dados a partir também da observação do verbo predicador/da sua natureza acional (dinâmica ou não) em relação com o padrão construcional de predicação configurado quanto à voz (passiva ou ativa) e, então, considerar também a construção de predicação verbal na voz ativa de relativa (des)focalização/opacificação dessa força indutora conforme a seguir:

Ex. 15 [SN<sub>Part2</sub> [V<sub>aux. de.voz passiva</sub> V<sub>auxiliado</sub>]<sub>predicador verbal complexo</sub> (por SN<sub>Part1 genérico</sub>)]<sub>predicação verbal com desfocalização/opacificação da força indutora</sub>

“Na sentença, o juiz Diógenes Luiz de Almeida Rodrigues afirmou que **foram criados** cargos comissionados para atribuições meramente técnicas e operacionais, como administrador regional, consultor jurídico especial do Procon, superintendente da Guarda Civil Metropolitana e assistente comunitário. (01 DE SETEMBRO DE 2017 <https://www.revistaforum.com.br/do-dcm-doria-nomeia-na-sua-equipe-prefeito-condenado-por-corrupcao/>)

Ex. 16 [Predicador verbal<sub>simples ou complexo</sub> -SE<sub>indeterminador de ParticipanteI</sub> SN<sub>Part2</sub>]<sub>predicação verbal com desfocalização/opacificação da força indutora</sub>

“Nas subcomissões que brotaram aqui e acolá, **criaram-se** cargos às dezenas, teimou-se em revisitar a morte de JK e Jango, descobriram-se...” ( 27 de fev de 2014 <https://oglobo.globo.com/cultura/ditadura-reloaded-11725856>)

“Projeto de Lei Complementar 12 de 06 de Junho de 2017, que dispõe sobre criação de vagas e cargos no Quadro de Pessoal da Prefeitura Municipal de Senador Amaral e dá outras providências: **criou-se** cargos na Secretaria de Educação;” (14 de julho de 2017 | Ano

<sup>29</sup> “No contexto de pares em alternância, o foco na semântica construcional e na compatibilidade semântica levanta várias questões: primeiro, quais são exatamente as diferenças semânticas (muitas vezes supostamente tênues) entre os membros desse par; segundo, quão produtiva é a ‘alternância’ no uso real, ou seja, quais verbos/substantivos ocorrem livremente em ambas as construções e quais têm fortes vieses/propensões em relação a uma delas; e terceiro, é plausível uma abordagem construcionista não derivada, dadas as respostas para as duas primeiras perguntas. Acreditamos que um método que amplia a noção de colocações distintas no contexto de nossa ‘collostructional analysis’ previamente proposta pode propiciar respostas para essas perguntas”.

10 | Edição 296 | Distribuição gratuita [http://www.jornalgazetadovale.com.br/wp-content/uploads/2017/07/gazetadovale\\_edicao296.pdf](http://www.jornalgazetadovale.com.br/wp-content/uploads/2017/07/gazetadovale_edicao296.pdf)

Ex. 17 [ (SN<sub>Part1 genérico</sub>) **Predicador verbal**<sub>simples ou complexo</sub> SN<sub>Part2</sub>] predicação verbal com relativa (des)focalização/opacificação da força indutora

“Vagas PCD / PNE???” **Criaram** uma nova função profissional????” (título de um texto publicado em 1 de fevereiro de 2016; <https://pt.linkedin.com/pulse/vagas-pcd-pne-criaram-uma-nova-fun%C3%A7%C3%A3o-profissional-claudemir>)

Nos três exemplos, encontram-se predicações cujo predicador é uma forma do lexema “*criar*”. Neles, vemos materializados três padrões construcionais diferentes de predicação verbal com (des)focalização/opacificação de força indutora, que se atualizam no discurso/texto no intuito de, em alguma medida, “tirar de cena” o participante (força indutora) envolvido na predicação da proposição/estado de coisas conceptualizada(o):

(i) um padrão construcional de predicação em que o verbo predicador se materialize na forma de uma locução verbal que sinalize o caráter passivo do participante SN<sub>Participante2</sub>, bem como a opacificação do participante 1 (força indutora), mesmo que este seja expresso por forma pronominal de valor indefinido ou por expressão de conteúdo genérico – [SN<sub>Part2</sub> [V<sub>aux. de.voz passiva</sub> V<sub>auxiliado</sub>] predicador verbal complexo (por SN<sub>Part1 genérico</sub>)] predicação verbal com desfocalização/opacificação da força indutora – (ex. 15) “foram criados cargos comissionados”;

(ii) um padrão construcional de predicação em que o verbo predicador se materialize na forma de verbo simples (flexionado ou não segundo o número do SN<sub>Participante2</sub>) ou complexo (com verbo auxiliar também flexionado ou não) acompanhada de pronome clítico SE (apassivador/indeterminador do SN<sub>Participante2</sub>) e, com essa formulação, sinalize o caráter passivo do participante SN<sub>Participante2</sub>, bem como a opacificação do participante 1 (força indutora), e, então, suspenda o SN<sub>Participante1</sub> – [Predicador verbal<sub>simples ou complexo</sub> -SE<sub>indeterminador de Participante1</sub> SN<sub>Part2</sub>] predicação verbal com desfocalização/opacificação da força indutora – (ex. 16) “criaram-se cargos” e “criou-se cargos”<sup>30</sup>;

(iii) um padrão construcional de predicação em que o verbo predicador se materialize na forma de verbo simples (flexionado na terceira pessoa do plural) ou complexo (com verbo auxiliar também flexionado na terceira pessoa do plural) acompanhado ou não de SN<sub>Participante1</sub>, expresso por forma pronominal de valor indefinido ou por expressão similar de conteúdo genérico e, com essa formulação, sinalize a opacificação do participante (força indutora) SN<sub>Participante1</sub>, embora, agora, com (algum) foco, já que a atenção se volta para o

<sup>30</sup> e, ainda, embora não estejam no exemplo, constructos como “podem-se criar cargos” ou “pode-se criar cargos”.

evento em si, para a ação/causação/indução (de “criar”) e o indutor/trajetor é perfilado – [ (SN<sub>Part1</sub> genérico) Predicador verbal<sub>simples ou complexo</sub> SN<sub>Part2</sub>] predicação verbal com relativa desfocalização/opacificação da força indutora

Podemos representar o espaço de variação construcional na rede construcional com base no mapeamento geral da relação de alinhamento funcional desses padrões conforme a seguir:

**Esquema 2:** Representação de aloconstruções na metaconstrução, configurada segundo valores mais produtivos dos atributos dos polos formal-funcional/do sentido.

		Predicação verbal com desfocalização/opacificação da força indutora do estado de coisas (dinâmicos ou não)		
		Aloconstrução 01	Aloconstrução 02	Aloconstrução 03
<b>Metaconstrução</b>		[SN <sub>Part2</sub> [V <sub>aux.voz</sub> V <sub>auxiliado</sub> ]predicador verbal complexo (por SN <sub>Part1</sub> genérico) ] <i>foram criados</i>	[Predicador verbal <sub>simples/complexo</sub> SE <sub>indeterminador de Participante1</sub> SN <sub>Part2</sub> ] <i>criaram-se / criou-se</i>	[Predicador verbal <sub>simples ou complexo</sub> SN <sub>Part2</sub> ] <i>criaram</i>
		<b>ATRIBUTOS – POLO DA FORMA</b>		
	Posição do predicador na oração (+/- inicial)	Posição não-inicial	Posição não-inicial	Posição não-inicial
	Posição do SNparticipante 1	quando não omitido, posposto ao predicador verbal	posição suspensa	quando não omitido, anteposto ao predicador verbal
	Posição do SNparticipante 2	anteposto ao predicador verbal	posposto ao predicador verbal	posposto ao predicador verbal
		<b>ATRIBUTOS – POLO DO SENTIDO</b>		
	Perfilamento do SNparticipante 1	normalmente, sem perfilamento; quando não omitido, perfilado indiretamente por SP contendo SN pleno genérico/de conteúdo indefinido, sem foco	perfilado indiretamente por pronome SE, sem foco	perfilado indiretamente por morfema gramatical de 3ª pessoa do plural no predicador (e, às vezes, por SN pleno genérico/de conteúdo indefinido), sem foco
	Perfilamento do SNparticipante 2	perfilado com foco	perfilado com foco	perfilado com algum foco, especialmente quando o SNparticipante 1 é omitido

É claro que os fatores dos atributos apreendidos nos padrões construcionais sugeridos<sup>31</sup> não explicam diretamente todas as diferenças ou similaridades entre

<sup>31</sup> Com base em investigação sobre a temática em curso no Projeto PREDICAR – Formação e

as construções. Reforça essa ideia o fato de que ainda vai atuar fortemente no recurso a um ou a outro padrão construcional para a efetivação da (des)focalização/opacificação do participante força indutora o *paradigma discursivo* (idealizado por WIEDEMER; MACHADO VIEIRA, 2018a, com inspiração em “padrão discursivo/*discursive pattern*”<sup>32</sup> em LEINO; ÖSTMAN, 2005).

Muito resumidamente, podemos relacionar, com base em resultados de pesquisa (em andamento no Projeto PREDICAR), o segundo padrão construcional, [Predicador verbal<sub>simples ou complexo</sub> -SE<sub>indeterminador de Participante</sub> SN<sub>Part2</sub>] predicação verbal com desfocalização/opacificação da força indutora, como mais fortemente atraído para esse tipo de predicação quando a conceptualização de estados de coisas se dá no domínio acadêmico e, em geral, em certos gêneros relativamente estáveis como artigos científicos, teses e dissertações.

É necessário investir na relação entre contextualidade e a atualização de (certos) padrões construcionais e (certos) lexemas na construção do sentido. Afinal, há atributos discursivos e pragmáticos que vão influir no maior ou menor acionamento (ou até não acionamento) de certos padrões e até de funcionalidades/significados deles. De acordo com Mackiewicz e Riley (2003, p. 84):

Briefly defined, pragmatics is the branch of linguistics concerned with how language use and interpretation are affected by specific contexts. Context includes variables such as the identity of the speaker and listener (for example, their relative social status), the speaker’s intent in producing a particular utterance (for example, whether the speaker is trying to inform or to persuade), and the linguistic conventions associated with particular intents (for example, the various ways in which requests are typically phrased).<sup>33</sup>

---

expressão de predicados complexos: estabilidade, variação e mudança construcional, cujos primeiros dados relativos ao Português de Portugal já resultaram no artigo: MACHADO VIEIRA, SANTOS; KROPF (2019).

<sup>32</sup> Leino e Östman (2005) valem-se da noção de “padrão/paradigma discursivo”. Segundo os autores, “A discourse pattern is the cognitive correlate of the linguistically defined *text type*, and the socioculturally defined *genre*. Understanding of text and discourse takes place primarily in terms of discourse patterns” (p. 200). “Padrão/paradigma discursivo” alinha-se a enquadre – Semântica de Frame – e é concebido para ser invocado por construção.

<sup>33</sup> “Resumidamente definida, a pragmática é o ramo da linguística preocupado com o modo como o uso e a interpretação da linguagem são afetados por contextos específicos. Contexto inclui variáveis, tais como a identidade do falante e do ouvinte (por exemplo, seu status social relativo), a intenção do falante em produzir um enunciado particular (por exemplo, se o falante está tentando informar ou persuadir), e as convenções linguísticas associadas a intenções particulares (por exemplo, as várias formas em que os pedidos são tipicamente fraseados”.

Goldberg (2016) já chama a atenção para a relação entre contextualidade e composicionalidade, destacando que o significado/a funcionalidade de expressões pode reconfigurar-se a depender do contexto em que elas se inserem/instauram.

Por agora, vislumbramos o paradigma teórico-metodológico aqui sinteticamente esboçado para o tratamento da variação de aloconstruções associadas em metaconstrução e (co)lexemas em *slot* construcional, cuja heurística, naturalmente, ainda carece de aprofundamento, principalmente, no que concerne aos métodos de análise de colocação construcional. Para isso, nossos esforços e próximos passos se concentram no aprofundamento do estudo da associação, por um lado, entre construções gramaticais e unidades lexicais<sup>34</sup> e, por outro, entre lexemas, construções gramaticais e paradigma discursivo, com a seguinte agenda de tópicos de pesquisa:

- a) Investigar a relação entre os itens lexicais e a construção gramatical – modelar o cálculo para  $n$  palavras e o quão fortemente essas palavras são atraídas para um *slot* em uma construção;
- b) Averiguar o contraste entre duas ou mais construções no que diz respeito aos itens lexicais que ocorrem nelas, ou seja, comparar construções que são aproximadamente sinônimas – modelar o cálculo para  $n$  palavras e o quão fortemente essas palavras são atraídas para duas ou mais construções funcionalmente semelhantes e, assim, traçar o perfil de relação lexemas-construções;
- c) Examinar o grau de dependência entre itens lexicais que ocupam dois *slots* diferentes dentro de uma mesma construção – modelar o cálculo para  $n$  palavras em um *slot* de uma construção e o quão fortemente unidades lexicais são atraídas para um *slot* por conta da influência de unidades em outro *slot* da mesma construção;
- d) Pesquisar a relação entre lexemas, construções e paradigma discursivo – modelar o cálculo de  $n$  palavras e/ou padrões construcionais e o quão fortemente esses elementos são atraídos para um paradigma discursivo e não outro.

Entre as vias de tratamento estatístico nos materiais reunidos, destacamos uma que permite lidar com a possibilidade de significados mais ou menos similares de certos lexemas (*colexemas*): a verificação de colocações lexicais (*collocational*

---

<sup>34</sup> Conforme empreendido por Ferreira (2019).

*analysis*). Nesse sentido, *corpora* podem ser explorados para tratamento estatístico de lexemas, em *slots* construcionais. E a avaliação de similaridade entre unidades lexicais pode enveredar pela escala de *prototipicidade* e/ou escala de *semelhança familiar* como referência para a generalização sobre as unidades lexicais que podem preencher o *slot* construcional. No primeiro caso, o exame leva à delimitação de: membros exemplares/mais centrais da categoria (com similaridade à configuração do protótipo implicada pelas condições/propriedades para preenchimento do *slot* construcional), membros intermediários e membros periféricos (efeitos do protótipo, com poucos traços em comum com a configuração característica do núcleo da categoria/do protótipo). No segundo, o exame leva ao mapeamento de zonas de sobreposição/mesclagem (metafórica ou não) na avaliação estatística por parte do falante dos atributos dos membros da categoria.

## O DESAFIO DE OPERACIONALIZAR UMA ANÁLISE DE VARIAÇÃO CONSTRUCIONAL

Traçar um desenho metodológico de partida para lidar com as possibilidades aqui aventadas num modelo socioconstrucionista de pesquisa não faria jus à dinamicidade e à plasticidade da rede de construções que configura um sistema linguístico. De todo modo, começamos a pensar a respeito de como direcionar nossas próprias investigações nesse modelo e vamos, então, esboçar um percurso que até então nos ocorreu.

Nesse percurso, o tratamento da variação construcional pode alcançar duas grandes vertentes: (i) uma que lide com o caráter da variação como atrelado a sentido referencial ou sócio-discursivo-pragmático; e (ii) uma que lide, efetivamente, com o caráter polifatorial (com propriedades de atributos) e polidimensional (com atributos diversos) de ambas as faces (forma-função/significação). E, então, pode chegar a operar, mediante introspecção e empiria, com as noções de prototipicidade, similaridade configuracional, semelhança simbólica e com processos cognitivos gerais (analogia, (re)combinação ou neoanálise), que envolvem gradualidade e se sujeitam a reconfiguração a cada atividade discursiva situada.

No percurso, consideramos, ainda, a possibilidade de compatibilização de abordagens metodológicas num estudo socioconstrucionista: método observacional/etnográfico, método experimental; método introspectivo e método empírico. Afinal, conforme a observação de Talmy (2014, p. xi), nenhuma metodologia deve ser super ou subestimada no processo de investigação, uma vez que cada uma lhe confere um perfil e tem capacidades e limitações.

Para desenvolver um estudo socioconstrucionista, cogitamos de uma configuração de condições, materiais, processos e procedimentos de trabalho analítico, entre os quais estão os que passamos a listar, sem, neste capítulo, entrar no detalhamento<sup>35</sup>. Com isso, prevemos procedimentos<sup>36</sup> como:

- Desenho de uma variável de construções e/ou (co)lexemas para análise, com base numa configuração da face de significação/função (a) que não se limite a funcionamento semântico nem à condição de identidade semântica/sinonímia (em sentido restrito) e (b) que contemple prototipicidade e relativa estabilidade na comparabilidade/associação entre construções ou (co)lexemas rotinizada e/ou mentalmente acessada por membros de uma comunidade num domínio discursivo-pragmático, social, histórico-cultural e/ou cognitivo, na definição de variantes/alternativas linguísticas. Projeção de problemas e hipóteses/previsões teórico-explicativas sobre um fenômeno que se suspeite/considere variável para encaminhar condições para a configuração de amostras de materiais e/ou processos a pesquisar<sup>37</sup>.
- Relação entre condições e escolhas relativas a, por exemplo: (a) conhecimento ou não de informantes, (b) materialidade do que estará sob observação<sup>38</sup>, (c) suporte dessas produções em relação com um domínio de comunicação<sup>39</sup>, (d) técnicas e/ou dispositivos/ferramentas para reunião de produções, materiais, bem como para seu tratamento estatístico (e) nível de

---

<sup>35</sup> Este ficará para outra oportunidade. A previsão aqui feita será explorada em estudo socioconstrucionista a ser efetivamente desenvolvido. A experiência deste promoverá a revisão do que aqui apenas em tese se expõe.

<sup>36</sup> Listamos procedimentos, sem, pelo menos no momento, preocupação com uma ordenação cronológica destes, até porque prevemos, no desenrolar da investigação, avanços e retornos.

<sup>37</sup> Nesse passo a passo naturalmente estão previstas também etapas que se aplicam a qualquer estudo/projeto linguístico.

<sup>38</sup> Produções orais, produções escritas, produções não-verbais, por exemplo. Neste terceiro caso, ocorreram-nos, por exemplo, linguagem corporal, expressões faciais, movimento ocular, explorados em pesquisa experimental do tipo *on-line* (que propicia medidas durante o processamento cognitivo do estímulo).

<sup>39</sup> Entre outros exemplos, ocorrem-nos estes: conversas quotidianas, inquéritos (mais ou menos controlados) gravados/filmados do tipo diálogo entre informantes ou entre documentador e informante ou do tipo elocução formal, conversas ou entrevistas em rádio, televisão e sítios acessados via internet; peças, roteiros de cinema, cartas, textos em jornais ou revistas (acadêmicos ou não); formulários para preenchimento, seleção e/ou avaliação (escalar) de respostas a estímulos.

representatividade das produções, dos materiais, (f) nível e natureza<sup>40</sup> de acesso às propriedades funcionais/de significação, entre outras.

- Exame preliminar da relação multifatorial e multidimensional dos atributos das faces forma e função/significação com base numa amostra prévia de constructos que instanciem as variantes construcionais investigadas e reflexão sobre processos cognitivos e relações envolvidas.
- Projeção do fenómeno variável em análise na estrutura cognitiva de uma língua, na sua rede de construções. Em outras palavras, representação formal e funcional das variantes construcionais em microconstruções e (sub) esquemas construcionais e de suas relações (de herança e extensão), bem como da relação de instanciação de tais padrões construcionais mentais/abstratos nos constructos detectados em produções linguísticas ou nos constructos sob processamento: (a) tanto no sentido do constructo/experiência de uso/processamento para a representação mental em padrões construcionais (microconstruções, mesoconstruções/subesquemas e macroconstruções/esquemas) quanto no sentido da construção (mais esquemática, passando por menos esquemáticas) até constructos/experiência de uso/processamento; (b) com base em parâmetros construcionais como esquematicidade, produtividade (*type* e *token*), composicionalidade e contextualidade.
- Observação, em tempo real e/ou aparente, de registros reunidos no uso e/ou via técnicas de pesquisa experimental para a detecção de propriedades relativas aos atributos forma e função/significação fundamentada em experiências de uso e/ou processamento (LABOV, 1994, 2001).
- Comparação desses registros com dados de natureza privada e/ou analiticamente manipulados, simulados ou controlados (com o recurso, por exemplo, de exame amostral via ferramenta de busca na internet e via dispositivos digitais referentes a linhas de concordância para reunião “rápida” de dados que se tencione averiguar), para testagem dos pareamentos identificados, da força de atração/restricção de uma construção, da força de coerção de uma construção, de *insights* sobre fenómenos novos.

---

<sup>40</sup> Por meio do vocábulo nível, referimo-nos à relação das produções reunidas com outras, até de dimensão mais complexa. Por natureza, referimo-nos às condições de (i) estudo de produtos/registros e/ou processos que os geram e (ii) reunião de evidências replicáveis noutros ambientes (manipulados, simulados e/ou controlados).

- Ainda com relação ao tratamento qualitativo de materiais e processos, mapeamento de variáveis/conjunto de fatores<sup>41</sup>: generalizações que viabilizem a configuração de variáveis para controle, para testagem e/ou para análise multivariada do fenômeno variável.
- Tratamento estatístico, principalmente, no que concerne à investigação da associação entre construções gramaticais e itens lexicais, conforme já destacamos, na seção anterior, ao lidar com variação de (co)lexemas em *slot* construcional. Análise de colocação e *collostructional analysis*<sup>42</sup>.
- Reunião, descrição, interpretação e explicação, baseada em reflexão/abstração e em observação a partir de experiências de uso/processamento, de resultados angariados nos procedimentos analíticos (de produtos e processos), com chance de refinamento de algum procedimento adotado ou reconfiguração de alguma representação categorial considerada no desenrolar da investigação.

## DESTAQUES FINAIS: REPERCUSSÃO DE ESTUDOS NUMA INTERFACE *VARIAÇÃO-(META)CONSTRUÇÃO*

Em linhas gerais, tais estudos podem abrir caminho para que o fenômeno de variação venha a se consolidar como uma das preocupações centrais entre as generalizações sobre a Gramática de Construções de uma língua configurada a partir (i) da relação entre experiências de uso, percepção e/ou processamento e representações cognitivas que pareiam atributos de forma e função/significação em (sub)esquemas construcionais, (ii) da concepção de articulação entre léxico e estrutura na gramática e (iii) da concepção desta como inventário de padrões delineado coletivamente a partir de experiências individuais.

E esses estudos podem contribuir para que tal fenômeno venha a ser mapeado nas generalizações sociolinguísticas com base no exame não só de propriedades relativas a atributos discursivos (como, por exemplo, grau de conexão discursiva de referente, plano discursivo – figura e fundo) ou pragmáticos (como, por exemplo, o estatuto informacional de referente definido, entre outras coisas, em

---

<sup>41</sup> A análise da variável dependente (de aloconstruções ou (co)lexemas) será mapeada com base na coatuação de atributos formais e funcionais (captados ou projetados) administrados pelo pesquisador na observação de registros de uso ou nos estímulos eleitos/configurados para as tarefas em pesquisa experimental.

<sup>42</sup> Klavan (2012), por exemplo, ao tratar de sinonímia gramatical, alia análise multivariada e *collostructional analysis*.

função de conhecimento de mundo (mais ou menos) partilhado, conhecimento pragmático de (inter)locutor e/ou de projeção sobre o conhecimento pragmático de interlocutor), mas também de propriedades relativas a: (i) atributos cognitivos (como, por exemplo, o grau de atenção do interlocutor a uma entidade/evento-referente, ao espaço temporal, geográfico ou mental de localização desta ou às imediações da interação/comunicação, a perspectiva/o ponto de vista de conceptualização de um estado de coisas/de uma entidade pelo enunciador, o grau de (inter)subjetividade na representação do estado de coisas conceptualizado, o grau de proeminência semântica na codificação de um conteúdo conceptual por mais de uma expressão linguística); e (ii) atributos sociais e culturais que passam a ser delineados/recontextualizados por orientação multidimensional (em termos sociohistórico-culturais, socioculturais, sociointeracionais) e apreendidos mediante o exame de registros de uso/produção, processamento, percepção, avaliação, crença, atitude e reação (afeto/emoção).

Aos atributos considerados no sentido do mapeamento cognitivo da variação com base na experiência, relacionam-se, então, observações quanto a representações psicológicas/mentais individuais (significados relativos à cognição individual) e representações psicológicas/mentais rotinizadas/convencionalizadas numa comunidade (significados relativos a uma cognição social, histórico-cultural, socialmente interativa). E ao tratamento escalar das categorias sob análise, soma-se a concepção de categorização de construções baseada em rede de relações (de herança, extensão e instanciação) e em organização radial das unidades ((sub)construções e dos constructos) a uma construção nuclear/protótipo numa estruturação categorial que abarca membros menos e mais exemplares (em outras palavras, membros periféricos e membros centrais), a fim de lidar, por exemplo, com domínios de ambiguidade referencial e de mesclagem conceptual (de correspondência/integração analógica entre membros pertencentes a domínios diferentes), entre outras operações cognitivas.

## **DISCUSSÃO FINAL**

Aqui procuramos corporificar parte do que já concebemos como respostas a perguntas formuladas a partir de Machado Vieira (2016), entre as quais: Qual é o lugar da variação na Gramática de Construções? Variação e alternância são termos diferentes? Que configuração pode ter variação por similaridade? Como se operacionaliza a representação desse tipo de fenômeno linguístico na rede construcional? Na relação forma-função, cogita-se da noção de protótipo ou semelhança de família? Recorre-se a um programa de interface (Sociolinguística

Variacionista e Gramática de Construções) ou formula-se uma heurística própria para a investigação desse tipo de fenômeno?

Em linhas gerais, nosso entendimento é o de que *variação* e *alternância* se prestam a designar o mesmo fenômeno, que, na Gramática de Construções, se instaura, via analogia, pelo menos em razão de duas possibilidades: de mais de uma forma se compatibilizar a *slots* construcionais ou pela possibilidade de construções distintas desenvolverem *links* associativos de similaridade/semelhança simbólica. E, em ambas as relações, cogitamos da noção de protótipo, uma vez que, estando a força de atração/repulsa em jogo, há (co)lexemas e padrões construcionais mais ou menos produtivamente acionados. Para lidar com generalizações dessa ordem, entendemos que uma heurística socioconstrucionista valendo-se de orientações e conceitos oriundos da Sociolinguística Variacionista e da Gramática de Construções merece ser testada: conjugando-se, inclusive, análise multivariada e análise construcional; considerando-se não só situações de relação entre variação e mudança, mas também situações de variação estável, de convivência de padrões construcionais/lexemas variantes; recorrendo-se, na representação, aos constructos teóricos da similaridade configuracional ou semelhança simbólica e da metaconstrução (espaço, na rede construcional de uma língua, subespecificado, para dar lugar à neutralização de diferenças entre lexemas/padrões construcionais que se entrelacem, a partir da experiência, por *links* associativos, por conta de proximidade, similaridade ou sinonímia, promovidos pelos usuários da língua).

Isso posto, passamos, então, a outra rodada de discussões, em que questões como as seguintes passam a impor-se: Como representar a sobreposição de feixes de propriedades de atributos de padrões construcionais/lexemas, que se (re)configuram estatisticamente, na rede construcional? E como fazê-lo, considerando as diversas configurações decorrentes de diferentes práticas discursivas e redes sociais de que indivíduos e grupos de falantes participam e que, portanto, experienciam? Como articular isso com expectativas e construções da ordem do domínio discursivo-pragmático e histórico-cultural em que tais configurações se instalam? Enfim, como podemos acessar e apreender a complexidade, a dinamicidade e a plasticidade de relações entre padrões construcionais/lexemas que indivíduos e comunidades podem estocar entre as generalizações que fazem de sua língua vivenciando fartas experiências linguísticas? Como reunir e tratar rigorosamente amostras representativas dessa realidade multifacetada para configurar generalizações linguísticas que estão/são estocadas na mente?

## REFERÊNCIAS

ALVES, Olívia. M. *Estudo sociofuncionalista da alternância entre predicadores pronominais simples e predicadores complexos*. Dissertação de Mestrado (Língua Portuguesa). UFRJ, 2011.

BOYD, Jeremy K.; GOLDBERG, Adele E. Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language*, 87, p. 55– 83, 2011.

CAMACHO, Roberto G. Construções passiva e impessoal: distinções funcionais. *Alfa*, São Paulo, 44:215-233, 2000.

CAPPELLE, Bert. Particle placement and the case for “allostructions”. *Constructions. Special Volume 1*, p. 1-28, 2006.

ESTEVES, Giselle Aparecida T. *Construções com DAR + sintagma nominal: a gramaticalização desse verbo e a alternância entre perífrases verbo-nominais e predicadores simples*. Dissertação (Mestrado em Letras Vernáculas) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

GOLDBERG, Adele E. *Constructions: a construction grammar approach to argument structure*. Chicago: The University of Chicago Press, 1995.

GOLDBERG, Adele E; BOYD, Jeremy K. A-adjectives, statistical preemption, and the evidence: Reply to Yang (2015). *Language*, vol. 91, n. 4, p. e184-e197, 2015.

GOLDBERG, Adele. Compositionality. In.: RIEMER, N. (Ed.) *Semantics Handbook*. Routledge, 2016, p. 419-430.

GRIES, Stefan Th.; STEFANOWITSCH, Anatol. Extending collocation analysis: a corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics*, 9:1, p. 97-129, 2004.

HILPERT, Martin. Collostrucional analysis: measuring associations between constructions and lexical elements. In.: GLYNN, D.; ROBINSON, J. A. (Eds.). *Corpus Methods for Semantics: Quantitative studies in polysemy and synonymy*. [Human Cognitive Processing, 43], p. 391-404, 2014.

KATO, Mary A. Português brasileiro falado: aquisição em contexto de mudança linguística. In: DUARTE, Inês; LEIRIA, Isabel. (Org.) *Congresso Internacional sobre o Português: actas*. Lisboa: A.P.L e Edições Colibri, 1996.

KLAVAN, Jane. *Evidence in linguistics: Corpus-linguistic and experimental methods for studying grammatical synonymy*. (Dissertationes Linguisticae Universitatis Tartuensis). Tartu: University of Tartu Press. 2012.

LABOV, William. *Principles of Linguistic Change: Internal Factors*. Oxford: Wiley-Blackwell, 1994.

LABOV, William. *Principles of Linguistic Change: Social Factors*. Oxford: Wiley-Blackwell, 2001.

LABOV, William. Some sociolinguistic principles. In: PAULSTON, C. B.; TUCKER, G. R. (Org.). *Sociolinguistics: the essential readings*. Oxford: Blackwell, 2003. p. 235-250.

LABOV, William. *Principles of Linguistic Change: Cognitive and Cultural factors*. Oxford: Wiley-Blackwell, 2010.

LEINO, Leino.; ÖSTMAN, Jan-Ola. Constructions and variability. In.: FRIED, Mirjam.; BOAS, Hans C. *Grammatical Constructions: back to the roots*. Amsterdam/Philadelphia: John Benjamins Publishing Company, p. 192-193, 2005.

MACHADO VIEIRA, Marcia S. Variação e mudança na descrição construcional: complexos verbo-nominais. *Revista Linguística*. / . Volume Especial, 2016, p. 152-170.

MACHADO VIEIRA, Marcia S.; WIEDEMER, Marcos L. A variação no modelo construcionista da linguística funcional-cognitiva. In: BRESCANCINI, Claudia Regina; MONARETTO, Valéria. *E-book do Encontro Intermediário do GT de Sociolinguística da ANPOLL*, 2017. (no prelo).

MACHADO VIEIRA, Marcia dos Santos; SANTOS, Júlia Lessa dos; KROPF, Morgana Pinheiro Albuquerque. Variação construcional por analogia: padrões construcionais de predicação verbal na voz passiva. *Soletras*, n. 37, p. 154-178, 2019.

PAVÃO, Bruna G.; MACHADO VIEIRA, Marcia S. Predicações com os verbos relacionais ser e estar. *Revista Diadorim*. Volume 14, p. 34-52, 2013.

PEREK, Florent. *Argument Structure in Usage-Based Construction Grammar: experimental and corpus-based perspectives*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2015.

ROBERT, Stéphane. Words and their meanings: principles of variation and stabilization. VANHOVE, Martine. *From polysemy to semantic change: towards a typology of lexical semantic associations*. John Benjamins, 2008, p. 55-92.

SILVA, Amanda; FONTENLOS, Clarissa; JUSTEN, Renata. Que tendências *se pode(m) encontrar* em textos escritos brasileiros? (Comunicação apresentada durante a 39ª Jornada Giulio Massarani de Iniciação Científica, Tecnológica, Artística e Cultural da UFRJ), 2017.

STEELS, Luc. Fluid Construction Grammar. In: HOFFMAN, Thomas; TROUSDALE, Graeme. *The Oxford Handbook of Construction Grammar*. New York, Oxford University Press, 2013, p. 153-167.

STEFANOWITSCH, Anatol; GRIES, Stefan Th. The Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8 (2), p. 209-243, 2003.

TALMY, Leonard. Foreword. In: GONZALEZ-MARQUEZ, Mónica., MITTELBERG, Irene., COULSON, Seana; SPIVEY, Michael J. (eds.). *Methods in Cognitive Linguistics*. Amsterdam: John Benjamins, 2007. p. xi-xxi.

TRAUGOTT, Elizabeth C.; TROUSDALE, Graeme. *Constructionalization and Construction changes*. Great Britain: Oxford University Press, 2013.

VAN DE VELDE, Freek. Degeneracy: The maintenance of constructional networks. In: BOOGAART, Ronny; COLLEMAN, Timothy; RUTTEN, Gijsbert (Eds.) *Extending the scope of Construction Grammar*. Berlin: De Gruyter Mouton, p. 141-179, 2014.

WIEDEMER, Marcos L.; MACHADO VIEIRA, Marcia S. Sociolinguística e Gramática de Construções: o envelope da variação. In.: FRANCESCHINI, L. T.; LOREGIAN-PENKAL, L. *Sociolinguística: estudos de variação, mudança e atitudes linguísticas*. Guarapuava: UNICENTRO, 2018a, p. 41-77.

WIEDEMER, Marcos L.; MACHADO VIEIRA, Marcia S. Lexemas e construção: atração, coerção e variação. In. *Caderno Seminal*. Dossiê especial “A centralidade da variação e a gradualidade da mudança linguística na língua em uso”, p. 81-132, 2018b.

# VARIATIONIST SOCIOLINGUISTICS AND CONSTRUCTION GRAMMAR THE CHALLENGES AND THE PROSPECTS OF COMPATIBILIZATION

*Marcia dos Santos Machado Vieira (UFRJ)*

*Marcos Luiz Wiedemer (UERJ)*

In this chapter, we argue in favor of the description of the phenomenon of variation by similarity/synonymy in the context of Portuguese Construction Grammar through the compatibilizing of Sociolinguistics orientations and concepts to Usage-based Construction Grammar orientations and concepts. In order to collaborate to construct a socioconstructionist heuristic, we use two theoretical-methodological constructs: (i) variation by allostructions (CAPPELLE, 2006) in metaconstruction (or constructeme, according to PEREK, 2015); (ii) variation by symbolic similarity. We also associate the following concepts: (i) the variation and change theory restriction problem (WEINREICH; LABOV; HERZOG, 1968) to the Usage-based Construction Grammar restriction or statistical preemption notions; (ii) restrictions or conditioning factors to (different or similar) factors/values involved in the formal and functional attributes of variant (lexical or grammatical) constructional patterns/allostructions; (iii) independent variables and attributes. Moreover, we signal a forwarding for the development of the scientific research: by focusing on the use and/or processing-perception-subjective evaluation experiences, as well as by considering a network of constructional patterns configured upon introspection and empiricism, it develops statistical generalizations about associative links between the values/factors of

the attributes involved the formal and functional faces of the constructions put/apprehended, in use, as being in variation by similarity/synonym relation.

Construction Grammar conceptualizes constructions as nodes in a hierarchical network lower-level constructions display a lower level of schematicity than the ones that license them; and multiple inheritance links govern the flow of features from the higher-level nodes to the lower-level ones. The links that hold between constructions/constructional patterns at relatively the same level of schematicity and which can enable language change (VAN DE VELDE, 2014) have only recently been object of observation, due to the assumption that, if constructions/constructional patterns are sufficiently similar on the same level, speakers might link them up in their mental grammars. So, similarity too is a factor that impacts the dynamic constructional network of a language, as soon as, a construction enters a context typical of another, their increased proximity/similarity may create an associative link.

Generally speaking, our contribution to such object of observation lies (i) in the fact that, in Brazil, only recently the question of variation by similarity/synonym is getting some scholars' attention and specially only after the discussion our texts (since MACHADO VIEIRA, 2016) have mobilized/provided here and (ii) in the fact that we have already initiated the configuration of a scientific path that articulates corpus-based discourse analysis, sociolinguistics multivariate analysis and constructional collocation/collocational analysis<sup>1</sup> (somehow, similar to the one that KLAVAN, 2012, has also developed) to deal with connections between variant constructional patterns and with the (variant) collexemes attraction patterns to the same constructional slot or the (variant) lexemes coercion patterns to a constructional slot.

In this effort to show how the configuration of constructional variation can take central place between the generalizations of experience-based Construction Grammar, we particularly problematize here two of the possibilities of variation treatment already mentioned: (i) variation by analogy/associative/alignment variation of independent constructions and (ii) variation by matching (co)lexemes to a constructional slot. In order to illustrate variation by symbolic similarity or by metaconstruction description and also variation by collocation of (co)lexemes, we use data from some ongoing researches in the PREDICAR Project / UFRJ<sup>2</sup>.

---

<sup>1</sup> Na linha de STEFANOWITSCH; GRIES (2003), GRIES; STEFANOWITSCH (2004).

<sup>2</sup> By the way, it is worth adding that the phenomenon of constructional variation is one of the thematics of two completed researches in such project: TRAVASSOS (2019) e FERREIRA (2019).

Theoretically, we conceive language as a cognitive and social system of constructions: a network of symbolic units that parallel attributes related to linguistic form (prosodic, phonetic, morphological, syntactic, lexical ones) to attributes related to semantic, discursive, pragmatic, social and cognitive functionality/meaning. This system conceives language in the language's (mental/emotional) use and processing experiences and is shaped by them. Further it is subjected to general cognitive processes such as: *analogical thinking* (process of combining aspects of function/meaning and form), which may or may not give rise to analogization (mechanism of change that gives rise to a previously non-existent correspondence), and *parsing* (process of analysis that motivates/enhances analysis/recombination of productive unit(s) different from the current analysis), that may or may not give rise to neoanalysis (mechanism of change that gives rise to a new structure). And change implies the stage of social dissemination and conventionalization of a (individual) linguistic innovation in a community. In this context, the constructional network of a language contains associative links put forward due to the conventionalized alignment of (some) formal and/or functional attributes values involved in independent constructional patterns or (co)lexemes. In other words, it means that there is an area in such network in which these constructional patterns or (co)lexemes have their own independent existence and there is also an area in which they align by attribute(s) value(s) association, by similarity/synonym relation, made/perceived by language community and, then, they are in *constructional variation*. Due to this theoretical conception briefly exposed, we seek to answer the question of how to delineate the design of the linguistic *variation-(meta)construction* in the proposal of compatibilizing Sociolinguistics and Construction Grammar.

Our hypothesis is the following one: if constructions are stabilized regular symbolic units that license the constructs (concrete utterance tokens that have actually been produced/processing), but are also subject to instability due to factors internal to the constructions (the statistical values of the attributes, as soon as constructions are potential prototype categories with more central and more peripheral features/values of attributes for providing the production/processament of constructs/uses/tokens), due to factors that are external to their manifestation/update in constructs/uses, and due to factors which are external and/or internal to speakers/communities of speakers involved in real situations of linguistic experience and/or linguistic processing, consequently the role of variation in the generalizations of the Construction Grammar of a language must be considered central. Notwithstanding, the role of variation can also be put as peripheral. The configuration of this role will depend on the problem to which the research pro-

posal is inclined: for variation; or for variation and change, or for change. The proposal will orient all the choices in the development of the research. We are arguing here in favor of the necessity and the feasibility of generalizations about variation that result from a linguistic analysis that is attentive to statistically interchangeable uses, that considers analogical thinking as one of the cognitive processes and that bases its generalizations on the theoretical assumption that the network of constructions of a language licenses uses and reconfigures itself when affected by uses in different language communities and/or different communities of practice or by uses in individual practices that influences those communities. Such statistical generalizations on constructional variation are representations that are updated through both the speaker's and the community's experience with language in a lot of different social, discursive, pragmatic and interactional contexts embedded in a cognitive, social and historic linguistic architecture and that may be mentally apprehended and/or assessed as being either in a stabilized coexistence situation or in a competition situation.

The linguistic units subject to the process of variation which we, in this chapter, focus on may have one of these configurations: (i) they are units/constructions of the order of metaconstruction (a constructional network overlapping area of representation in which there is a (relative) neutralization of the differences of the constructional patterns in variation in favor of their formal and/or functional attributes values in common), that is, *allostructions* (named according to CAPPELLE, 2006); or (ii) they are units/constructions of the order of the constructional slot, that is, *(co)lexemes* (form-functional pairing representations associated to the same constructional slot). The constructional slots, depending on the level of schematicity of the construction they are in, will be subject to more or less filling possibilities and to the compatibilization of more or less exemplary/prototypical members of the category. Another possibility of analytical cutout is the relation of both configurations, by calculating the attraction/repulsion force of (co)lexeme regarding a constructional slot (in a construction, in all constructions that configure the metaconstruction), as well as the profile of constructions or of allostructions with which the (co)lexemes maintain a matching relationship (by force of attraction/repulsion or force of coercion) in the language system.

The envelope design of the varying constructional patterns in a metaconstruction is made of two or more (procedural or lexical) constructions that are associated by analogy (because of their configurational similarities or because some symbolic association attributed to them), thus are foreseen in variation and, then,

are socially routinized as well as cognitively stored as constructional alternatives (allostructions). Metaconstruction (*constructeme*, as named by Perek, 2015) is a theoretical construct that, left partially underspecified, captures the level of representation at which constructions/constructional patterns systematically in alternation are functionally equivalent. Allostructions specify exactly how such constructions differ (by which attribute values/properties, restrictions). Thus, the metaconstruction/allostruction representation has the potential to apprehend the tension between similarity and dissimilarity (ie, values, in constructions, linked to their attributes).

In order to exemplify the alignment of the *allostruction* and *metaconstruction* concepts with the concepts involved in the configuration of the envelope of linguistic variation (variants and dependent variable), we use, among other Portuguese research results, the ones of the research about the *discursive impersonalization verbal predication construction*, conducted by Silva, Fontenlos and Justen (2017): [<sub>Auxiliar</sub> V <sub>Direct Transitive</sub> V + <sub>Participant passivation/indeterminacy</sub> SE SN]. These construction license strategies for signaling a participant's opacification/defocusing by discursive impersonalization such as: [<sub>Singular/Plural auxiliar</sub> V <sub>Direct Transitive</sub> V + <sub>Participant passivation/indeterminacy</sub> SE <sub>Singular/Plural Non-agentive</sub> SN] and [<sub>Singular auxiliar</sub> V <sub>Direct Transitive</sub> V + <sub>Participant indeterminacy</sub> SE <sub>Singular/Plural Non-agentive</sub> SN]. The former captures the possibility of variation in verbal flexion according to singularSN or pluralSN and being SE conceptualized as either a participant passivation resource or an indeterminacy one. The recognition of the passivation resource depends on the (plural) verbal flexion (according to pluralSN). The latter represents the stability of the singular verbal flexion, being the SN either singular or plural and being SE conceptualized necessarily as a participant indeterminacy resource. We represent this variation in Portuguese by configurational similarity (where patterns reveal stable parts and similar attributes, even because of the subpart inheritance relationship to which they are submitted, and also different attributes). *Allostructions* then designate the variant constructional patterns themselves, while *metaconstruction* is the nomenclature used to capture the shared generalization domain (of the expression of states of affairs with the opacification/defocusing of its inducing force – agent, causer, force, experiencer), in which those independent patterns (with different verbal flexional features) relate and are configured in an abstract neutralization area.

In order to illustrate the alignment of the *(co)lexemes* and *constructional slot* concepts with the concepts involved in the configuration of the envelope of linguistic variation (variants and dependent variable), we resort to Alves' (2011)

results about the variation between pronominal simple verb forms (*preocupar-se*, to worry) and similar periphrastic verbal-nonverbal forms (*ter preocupação*, to be worried) in the constructional slot of the verbal predicate/slot in a verbal predication. According to such research, there are some verbs that, by force of attraction, tend to match in the verbal slot of the periphrastic forms and there are some verbs that, by force of coercion, can also appear (although less in statistical analysis) in some verbal-nonverbal periphrasis. The verb *ficar* (in a similar use as “to get”, “to be” in “to get disturbed”/to disturb) is among the first ones. Then, two or more lexical units/lexical constructions have the potential to develop the profile of alternatives in the constructional slot matching process by attraction (collexemes) or coercion (lexemes). Therefore, a slot in a construction can have (i) units with formal-functional similarity, that is, alternatives attracted by the categorical slot/collexemes and then more expected due to the filling conditions implied by it and, naturally, by inclinations of use observed in a community; and (ii) less in tune units with the categorical slot attributes/features and, then, not or less attracted by the slot/lexemes, which, by virtue of coercion exercised by the construction, are formally and functionally aligned with those most often matched in such slot. They, then, act as alternatives to these. The (co)lexemes can, then, turn out to be variants of a dependent variable/constructional slot, which has fill constraints.

Furthermore, this chapter suggests the possibility of bringing together various empirical methods, such as corpus analysis or experimental studies, and the introspective method, to approach language variation within the theoretical framework of combining Sociolinguistics and Usage-based Construction Grammar and the mapping of a language constructional network, where we assume that speakers and communities (re)organize their mental grammar on the basis of the available input via cognitive processes (including analogical thinking, that can give rise to associative links) which result in a (kind of “diasystematic”) constructional grammar which comprises both language-specific and independent constructions restricted to certain formal and functional features/attributes and constructions with features/attributes to some extent underspecified, besides the underspecified or less specified slots in some constructions. In order to develop a socioconstructionist research, we propose some aspects that deserve attention in the configuration of analytical conditions, sample designs, research processes and procedures, among which there is the delimitation of the design of linguistic variation assumed here.

## REFERENCES

- ALVES, Olívia. M. *Estudo sociofuncionalista da alternância entre predicadores pronominais simples e predicadores complexos*. Dissertação de Mestrado (Língua Portuguesa). UFRJ, 2011.
- CAPPELLE, Bert. Particle placement and the case for “allostructions”. *Constructions. Special Volume 1*, p. 1-28, 2006.
- FERREIRA, Bruna Góis Pavão. *Construção predicativa de mudança de estado e de propriedade com os verbos ficar, tornar-se e virar*. Tese de Doutorado. UFRJ, Faculdade de Letras, 2019. 139f.
- GRIES, Stefan Th.; STEFANOWITSCH, Anatol. Extending collostructional analysis: a corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics*, 9:1, p. 97-129, 2004.
- KLAVAN, Jane. *Evidence in linguistics: Corpus-linguistic and experimental methods for studying grammatical synonymy*. (Dissertationes Linguisticae Universitatis Tartuensis). Tartu: University of Tartu Press. 2012.
- MACHADO VIEIRA, Marcia S. Variação e mudança na descrição construcional: complexos verbo-nominais. *Revista Lingüística. /*. Volume Especial, 2016, p. 152-170.
- SILVA, Amanda; FONTENLOS, Clarissa; JUSTEN, Renata. Que tendências *se pode(m) encontrar* em textos escritos brasileiros? (Comunicação apresentada durante a 39ª Jornada Giulio Massarani de Iniciação Científica, Tecnológica, Artística e Cultural da UFRJ), 2017.
- STEFANOWITSCH, Anatol; GRIES, Stefan Th. The Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8 (2), p. 209-243, 2003.

TRAVASSOS, Pâmela Fagundes. Variação e mudança construcional: um olhar funcional-cognitivo sobre usos de construções com verbo-suporte DAR. Dissertação de Mestrado. UFRJ, Faculdade de Letras, 2019. 260f.

VAN DE VELDE, Freek. Degeneracy: The maintenance of constructional networks. In: BOOGAART, Ronny; COLLEMAN, Timothy; RUTTEN, Gijsbert (Eds.) *Extending the scope of Construction Grammar*. Berlin: De Gruyter Mouton, p. 141-179, 2014.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin I. Empirical foundations for a theory of language change. In *Directions for historical linguistics: A symposium*, ed. W.P. Lehmann e Yakov Malkiel, Austin-London, University of Texas Press, 1975 [1968]. pp.95-199.

# PERCEPÇÃO DA VARIAÇÃO DA CODA (S) NA COMUNIDADE DE FALA DO RIO DE JANEIRO ACESSANDO O SIGNIFICADO SOCIAL DA VARIANTE FRICATIVA POSTERIOR

*Marcelo Alexandre Silva Lopes de Melo (UFRJ)*  
*Christina Abreu Gomes (UFRJ/CNPq)*

### RESUMO

O presente trabalho apresenta os resultados para um teste de percepção que avalia o significado social das variantes da variável coda (s) aplicado a três grupos de participantes pertencentes a dois grupos sociais distintos na comunidade de fala do Rio Janeiro. O compartilhamento de padrões de avaliação social entre falantes é um componente importante do pertencimento dos falantes a uma mesma comunidade de fala (LABOV, 1972). Diferentes estudos, baseados em dados de produção, fornecem evidência de que a variante fricativa posterior (velar/glotal), como em me[h]mo, é estigmatizada em diversas variedades do Português Brasileiro. Os resultados do teste aplicado revelaram que a dinâmica sociolinguística de falantes de uma determinada comunidade de fala é complexa, uma vez que, conforme pôde ser observado, não só existem diferentes avaliações das variantes da coda (s) entre os diferentes grupos de participantes, como também as diferenças são graduais. Os resultados apontaram ainda que padrões de avaliação estão intimamente relacionados ao grau de acesso dos indivíduos dos diferentes grupos às instituições que contribuem para a manutenção e implementação dos padrões de prestígio.

## INTRODUÇÃO

Este trabalho apresenta os resultados de um teste de percepção das variantes da coda (s), aplicado a falantes de diferentes grupos sociais da comunidade de fala do Rio de Janeiro (MELO, 2017), com o objetivo de identificar o significado social de duas das variantes fricativas, pós-alveolar [ʃ/ʒ] e velar/glotal [x/χ, h/h̥], como em me[ʒ]mo e me[h̥]mo. Participaram do teste falantes de dois grupos sociais distintos: a) um grupo de jovens universitários (grupo UFRJ), falantes de classe média-média e média-baixa; b) dois grupos de adolescentes moradores de favelas, falantes de classe baixa, com diferentes graus de inserção social (grupos EJLA e Fiocruz). Um dos objetivos para a realização do teste de percepção foi observar se os indivíduos de uma mesma comunidade de fala compartilham as mesmas avaliações sobre diferentes variantes de uma mesma variável. O compartilhamento de padrões de avaliação social entre falantes é um componente importante do pertencimento dos falantes a uma mesma comunidade de fala (LABOV, 1972).

Diversos estudos trataram da variação da coda (s) em diferentes comunidades de fala (CANOVAS, 1991; CALLOU; MORAES, 1996; RONCARATI, 1999; MOTA, 2002; MACEDO, 2004; GRYNER E MACEDO, 2000; BRESCANCINI, 2006; SANTOS, 2009), sendo alguns específicos sobre a cidade do Rio de Janeiro (CALLOU e MARQUES, 1975; GUY, 1981; SCHERRE e MACEDO, 2000). Na comunidade de fala do Rio de Janeiro, a coda (s) pode ser realizada como uma fricativa alveolar, alveolopalatal, velar ou glotal – respectivamente, como em me[z]mo ~ me[ʒ]mo ~ me[χ]mo ~ me[h̥]mo – ou ainda não ser realizada, como em me[0]mo. Alguns dos trabalhos anteriormente mencionados apontam para a existência de estigma relacionado à realização das fricativas posteriores (velar ou glotal) em função da distribuição das variantes por escolaridade, uma vez que a variante velar/glotal tende a ocorrer em falantes com baixa escolaridade. No entanto, nenhum estudo de percepção foi realizado para confirmar tal estigma, motivo pelo qual o presente trabalho pode contribuir para que se verifique o significado social dessas variantes em função de diferentes perfis sociais dos participantes.

Serão apresentados a seguir, na primeira seção, os conceitos principais que embasam este estudo: considerações sobre os estudos de percepção da variação, sobre o conceito de comunidade de fala e sobre os resultados de estudos sobre a variação da coda (s) que apontam para o estigma atribuído à variante fricativa posterior. A seção dois apresenta a metodologia de trabalho no que diz respeito ao *design* do teste aplicado, participantes e tratamento de dados.

As seções três e quatro, respectivamente, tratam da análise dos resultados e considerações Finais.

## PERCEPÇÃO DA VARIAÇÃO

Diferentes estudos sobre percepção da variação têm procurado avançar aspectos sobre a organização cognitiva da variação (CLOPPER; CONREY; PISONI, 2005; HAY; WARREN; DRAGER, 2006; STAUM-CASASSANTO, 2010; SQUIRES, 2011). Os trabalhos sobre percepção, embora ainda em menor número se comparados aos trabalhos sobre produção da variação, constituem um conjunto considerável de estudos que analisam não só a percepção da indexação social das formas linguísticas em uso, mas também o impacto dessas informações no processamento da linguagem.

Sobre a avaliação de uma determinada forma linguística por falante da cidade de Nova York, Labov (2006[1966]) sustenta que tal avaliação ocorre de acordo com a maior ou menor adoção dessa forma por classes sociais mais altas, ou seja, quanto mais uma forma linguística for usada por falantes pertencentes à classe social mais elevada, maior será o prestígio conferido a essa mesma forma. Labov sustenta ainda que os falantes não têm consciência sobre o uso de variáveis fonológicas, sendo, por esse motivo, necessário realizar diferentes testes para que seja possível aferir as avaliações realizadas pelos falantes sobre essas variáveis. Assim, o autor postula três principais problemas que devem ser resolvidos para que as questões ligadas à avaliação das formas linguísticas pelos falantes possam ser mais bem entendidas: a) isolar as reações subjetivas a valores particulares de uma única variável; b) reduzir essas reações a uma medida quantificável; c) encontrar a estrutura global refletida no padrão das medidas resultantes (LABOV, 2006, p. 266).

Labov et al. (2011) afirmam que experimentos que rastreiam os correlatos cognitivos da variação linguística sempre estiveram ao lado de estudos de produção de fala. Ainda de acordo com Labov et al. (op. cit.), a técnica de estímulos pareados ou técnica dos “falsos pares” (*matched guise test*), uma ferramenta metodológica desenvolvida por Lambert et al. (1960), contribuiu sobremaneira para os estudos sobre percepção. Como reações dos falantes/ouvintes podem não refletir diretamente suas opiniões pessoais, a técnica de estímulos pareados permite investigar atitudes subjetivas dos falantes/ouvintes que não são diretamente aferíveis. Labov (2008, p. 176) argumenta que o trabalho de Lambert é importante porque mostra a existência de “um conjunto uniforme de atitudes frente à linguagem que são compartilhadas por quase todos os membros

da comunidade de fala, seja no uso de uma forma estigmatizada ou prestigiada de uma língua em questão”, sendo mais provável que as atitudes de uma pessoa frente a um determinado uso não emergem de “forma sistemática se a pessoa for questionada diretamente sobre os dialetos”. Em outras palavras, as atitudes dos ouvintes somente emergem de forma sistemática se eles forem colocados diante de dois conjuntos de possibilidades produzidas por um mesmo falante usando duas formas diferentes. Para tanto, o ouvinte que avalia não deve perceber que as diferentes formas foram produzidas pelo mesmo falante. Oushiro (2015) observa que a técnica de estímulos pareados permitiu demonstrar que diferentes variedades linguísticas têm influência na avaliação dos sujeitos em diferentes situações, as quais vão desde a escolha de um inquilino ou de um candidato a uma vaga de emprego à associação da natureza de um crime cometido a falantes de uma determinada variedade.

Diversos estudos sobre percepção (LAMBERT, 1960; CAMPBELL-KIBLER, 2010; LABOV, 2008, 2011; OUSHIRO, 2015) indicam que, em relação à avaliação das formas linguísticas, há uma recorrente associação entre as variantes produzidas e o julgamento dos falantes que delas fazem uso. O prestígio ou estigma associado a uma determinada variante está associado a diversas questões e conformações sociais: as variantes de menor prestígio são identificadas como sendo características de grupos de falantes que pertencem a comunidades estereotipadas, a classes sociais menos privilegiadas ou menos escolarizadas. A partir de um teste de avaliação da variável (-r) na cidade de São Paulo, Oushiro (2015, p. 318) argumenta que “línguas, variedades e variantes favorecidas pelas classes menos privilegiadas são percebidas mais negativamente quanto a atributos de status e valorizadas quanto a traços de solidariedade”. Labov et al. (2011) investigaram a percepção da variação linguística por meio de uma série de experimentos elaborados para medir a sensibilidade dos ouvintes à realização apical não-padrão da variável (ING, como em *preparing* ‘preparando’). Os participantes foram expostos a diferentes frequências da variante IN para avaliar o grau de adequação da leitura de sentenças por uma suposta candidata ao cargo de locutora de programa de notícias de TV. Os resultados obtidos levaram os autores a concluir que, quanto mais a forma reconhecida como menos prestigiada - no caso, /in/ - era ouvida em um conjunto de sentenças, mais mal avaliada era a candidata.

Os resultados dos estudos anteriormente mencionados sugerem que as avaliações que os falantes fazem sobre as formas linguísticas em uso são construídas a partir das diferentes experiências sociais desses mesmos falantes. Votre (2010, p. 52) argumenta que “o modo de comunicação das pessoas

desprovidas de prestígio econômico e social tende a ser coletivamente avaliado como estigmatizado”. O autor prossegue dizendo que as formas estigmatizadas são associadas com um “falar errado” ou “vício” por parte dos falantes mais escolarizados. Assim, capturar a(s) avaliação(ões) das formas linguísticas e que são realizadas por falantes de diferentes grupos sociais de uma mesma comunidade de fala pode contribuir para que a dinâmica sociolinguística desta comunidade de fala seja mais bem compreendida, de maneira a explicitar as diferentes relações que os indivíduos estabelecem em razão de suas experiências concretas de uso e de sua própria identidade sociolinguística.

## CONCEITO DE COMUNIDADE DE FALA

Labov (1972) propõe que todos os indivíduos de uma mesma comunidade de fala compartilham as mesmas avaliações sobre as formas linguísticas. De acordo com o conceito de comunidade de fala formulado inicialmente por Labov e ampliado posteriormente por Guy (2001), “uma comunidade de fala não pode ser concebida como um grupo de falantes que usam todas as mesmas formas; ela é mais bem definida com um grupo que compartilha as mesmas normas a respeito da língua” (LABOV, 2008, p. 188). Assim, os falantes de uma mesma comunidade de fala apenas replicam em maior ou menor grau os padrões linguísticos estabelecidos por seus pares, não sendo, portanto, autônomos em relação à comunidade à qual pertencem. Isto implica dizer que, apesar de um falante poder fazer uso de uma forma a qual ele mesmo avalia como estigmatizada, as noções de estigma e prestígio são compartilhadas por todos os falantes de uma mesma comunidade de fala. Portanto, o próprio Labov (2008, p. 188) observa que, em relação à produção do (r) em final de sílaba, uma vez que falantes mais velhos e mais jovens de Nova York apresentam padrões diferentes de avaliação da realização da coda, então pertencem a “comunidades de fala ligeiramente diferentes”. Nesse sentido, reconhecer que há comunidades de fala ligeiramente diferentes dentro de uma mesma comunidade implica dizer que os falantes de uma mesma comunidade nem sempre compartilham das mesmas avaliações sobre as formas linguísticas.

No entanto, há autores que argumentam que há a possibilidade de existirem diferentes avaliações dentro de uma mesma comunidade. Santa Ana e Parodi (1998) observam que esse conceito foi concebido de maneira a relacionar dados sociolinguísticos a uma teoria linguística formal, ressaltando questões linguísticas em detrimento de questões sociais e sem conseguir capturar as funções sociais dos falantes com seus interlocutores. O conceito de comunidade de

fala postulado por Labov, segundo Santa Ana e Parodi, articula “produção” e “avaliação” a uma unidade de medida – a “variável linguística”, sendo o resultado de tal articulação expresso por meio de uma regra variável, a fim de que se garanta a uniformidade linguística na diversidade encontrada entre os falantes. Santa Ana e Parodi entendem que a hierarquia social reflete a avaliação que os falantes fazem acerca do modo como as pessoas falam, mas essa avaliação nem sempre é compartilhada por todos os falantes. Ao estudarem um dialeto do espanhol falado em uma região isolada do México, os autores defendem que diferentes avaliações das formas linguísticas empregadas por uma comunidade podem ser realizadas a partir do grau de acesso dos falantes às instituições e às formas de prestígio.

Relacionando a questão do compartilhamento de avaliações em uma mesma comunidade de fala à variável em análise no presente estudo, é possível observar que alguns trabalhos apontam para uma avaliação negativa da realização da variante aspirada (fricativas posteriores) em diferentes variedades do PB. Carvalho (2000) sustenta que a variante glotal é a menos prestigiada entre os falantes de Belém (PA) e, por isso, tende a ser usada por falantes pertencentes à classe baixa, sendo a realização da variante de prestígio (palatal) mais frequente entre os informantes do sexo feminino. Gryner e Macedo (2000, p. 44) sustentam que, na região de Cordeiro, “as formas mais populares (aspiração e queda)” tendem a ser realizadas por falantes menos escolarizados, sendo a variante palatal predominantemente realizada por falantes com alto nível de escolaridade. Auler (1992, p. 51) sustenta que o fenômeno da aspiração do *s* pós-vocálico, na comunidade de fala da cidade do Rio de Janeiro, deve ser tratado como um “indicador social de uma fala menos comprometida com a norma culta” e não como “um marcador dialetal diatópico”. Isto porque a realização da aspiração seria “própria de indivíduos que não estão sujeitos a pressões sociais sobre o seu desempenho linguístico”, tendo em vista que, ainda de acordo com a autora, os indivíduos que constituíram os *corpora*, ao aumentarem o seu nível de escolaridade, diminuíram ou eliminaram a pronúncia aspirada. No entanto, as hipóteses aventadas nos estudos mencionados acerca da avaliação das variantes da coda (*s*) precisam ser verificadas por meio de teste de percepção que comprove a existência de um estigma relacionado à variante velar/glotal. As assunções são feitas com base em dados de produção em função da distribuição das variantes considerando características sociais dos falantes que mais fazem uso da variante aspirada ou posterior, isto é, de falantes menos escolarizados e pertencentes a camadas mais populares.

Em Melo (2012, 2017), a partir de dados de produção com falantes de diferentes classes sociais – falantes da classe média média e média baixa, além de falantes de classe baixa, moradores de favelas, com diferentes níveis de escolaridade e graus de inserção social –, observou-se uma mudança em progresso em direção à variante velar/glotal somente entre os falantes socialmente excluídos da classe baixa. Foi possível observar ainda que, em todas as amostras analisadas, independentemente da classe social ou grau de inserção social, a variante velar/glotal tende a ser realizada quando a coda (s) é seguida por uma consoante sonora, em posição final de palavra, em palavras não-monossilábicas, em sílabas átonas, sem status morfológico, em estilo informal, sendo ainda condicionada em razão do indivíduo e do item lexical. A taxa global de realização de fricativas posteriores é muito superior entre os falantes socialmente excluídos da classe baixa (30%), sem escolarização regular, do que entre falantes da classe média média e média baixa (5%) e falantes da classe baixa com maior inserção social (6%).

Este estudo adota a hipótese dos modelos baseados em exemplares (BYBEE, 2001, 2010; PIERREHUMBERT, 2003; DRAGER; KITLEY, 2016), segundo a qual a variação linguística indexada é representada no léxico. Como o detalhe fonético é estocado como parte da forma da palavra, a representação é atualizada por meio da experiência do falante com a língua em diferentes contextos sociais, discursivos e interacionais. Em Melo (2012, 2017), assumiu-se que as diferentes taxas de realização das fricativas posteriores para itens específicos indicam qual variante é predominante – ou central – em relação às demais na representação da palavra para os indivíduos (CONNINE; RANBOM; PATTERSON, 2008). Assim, percebe-se que, entre falantes socialmente excluídos da classe baixa, a fricativa posterior prevalecia em alguns itens, tais como *mesmo*, *desde*, *vários* e *nós*, enquanto entre os falantes da classe média e os falantes da classe baixa com maior inserção social, esses mesmos itens foram produzidos predominantemente com a variante pós-alveolar. Assim, como a organização dos itens que contêm a coda (s) é diferente, pelo menos em relação a algumas palavras, entre os três grupos, é fundamental investigar até que ponto falantes com o mesmo perfil social do estudo de produção compartilham os mesmos padrões de avaliação entre as duas variantes (pós-alveolares e posteriores).

## METODOLOGIA DA PESQUISA

A avaliação social das variantes da variável linguística estudada foi acessada por meio de teste de percepção que observou a associação entre duas variantes da variável coda (s) – glotal e pós-alveolar – a três diferentes ocupações profissionais, que podem ser situadas em uma escala com diferentes graus de prestígio na sociedade: *médica*, *auxiliar de enfermagem* e *faxineira*. A associação entre variante e profissão revela expectativas de pertencimento a um determinado grupo social da comunidade de fala. À profissão de médica, além da obrigatoriedade de conclusão de um curso superior, é atribuído um prestígio social maior. Já a profissão de faxineira não necessita de formação específica e tampouco escolaridade, além de ser, geralmente, ocupada por pessoas de classes sociais mais baixas. Quanto a profissão de técnica em enfermagem há necessidade de determinada formação – mesmo que não seja necessariamente de ensino superior – e geralmente é ocupada por pessoas de classes mais populares que tiveram algum acesso a determinadas instituições sociais. Dessa forma, a associação de uma determinada variante aos extremos da escala de profissões indica: a) uma avaliação negativa da variante no extremo relativo à profissão de faxineira; b) e uma avaliação positiva no extremo relativo à médica. O perfil de técnica em enfermagem situa as variantes no meio de um *continuum* que teria, em seus extremos, o perfil de faxineira (estigma) em uma ponta e, na ponta oposta, o perfil de médica (prestígio).

Como um dos objetivos do estudo de Melo (2017) foi o de conjugar os resultados obtidos por meio do teste de percepção a dados de produção para indivíduos dos mesmos grupos sociais – ou com, pelo menos, padrões sociais semelhantes –, os participantes selecionados para o teste de percepção têm o mesmo perfil social dos falantes que compõem as amostras de fala do estudo de produção das variáveis estudadas: subgrupo de falantes da amostra Censo 2000, amostra Fiocruz e amostra EJLA. Assim, 36 indivíduos de dois grupos sociais participaram do teste de percepção: a) 12 indivíduos de classe média e média baixa, os quais foram representados por jovens universitários (grupo UFRJ); b) 24 indivíduos de classe baixa, os quais foram representados por dois grupos de adolescentes moradores de favelas: um grupo de 12 adolescentes com alguma inserção social, escolarização regular, os quais, no momento da aplicação do teste, participavam de um curso para monitor de museu em uma grande instituição pública de ensino e pesquisa do país (grupo Fiocruz); um grupo de 12 adolescentes excluídos socialmente, com escolarização bastante irregular, e que, no momento da aplicação do teste, cumpriam medida socioeducativa de interna-

ção em uma unidade do estado (grupo EJLA). Do total dos 36 participantes, 24 eram homens e 12, mulheres. Como não foi obtida uma boa distribuição por sexo nos três grupos, essa variável não foi considerada na análise.

As sentenças do teste foram gravadas por uma mulher, falante da variedade carioca, e, portanto, apresentando as duas variantes em questão na sua produção espontânea, com nível universitário e idade entre 25-30 anos. Além de 12 sentenças distratoras, 24 sentenças foram elaboradas e organizadas em duas listas: uma lista com 12 sentenças que só continham apenas um item com uma das variantes da variável em questão e outra lista com 12 sentenças que continham dois itens. As sentenças foram controladas em função do tamanho, contexto seguinte à coda e distância entre as duas palavras com a coda (s) na lista dos estímulos com mais de uma variante na mesma sentença (Anexo 1).

Cada participante ouviu, no total, 18 sentenças: 12 sentenças com uma ou duas variantes da coda (s) e 06 sentenças distratoras. Assim, metade dos participantes de cada grupo foi exposta a sentenças em que a variante ocorria uma vez em cada sentença e a outra metade foi exposta a sentenças com duas ocorrências da mesma variante, totalizando 18 participantes em cada uma das duas condições do teste (ouvir uma única ocorrência da variante por sentença *vs.* ouvir duas ocorrências da mesma variante por sentença). Cada participante ouviu, em cada uma das duas condições, 6 sentenças com uma variante e 6 com a outra variante, sendo que nenhum participante foi exposto à mesma palavra com as duas variantes. Foi produzido um total de 432 respostas, 216 relativas à variante pós-alveolar e 216 relativas à variante glotal. O teste foi aplicado individualmente nas respectivas instituições em que cada um dos grupos de participantes foi contatado. Os estímulos foram apresentados utilizando o software TP (Teste/Treinamento de Percepção) – versão 3.1, com apresentação dos estímulos em ordem aleatória.

No momento de aplicação do teste, o participante era apresentado à tarefa pelo pesquisador através de uma situação relatada através de um texto que aparecia na tela do computador. O texto a seguir era lido pelo pesquisador, acompanhado pelo participante:

*Você está em um hospital público, aguardando por uma consulta médica. Enquanto aguarda ser atendido, você ouve uma mulher dizendo algumas frases. Após ouvir cada frase, diga se ela foi produzida por uma médica do hospital, por uma técnica de enfermagem do hospital ou por uma auxiliar de serviços gerais.*

Após o texto, o participante ouvia o estímulo. Cada sentença poderia ser ouvida até 03 (três) vezes pelo participante antes de indicar a resposta. Após ouvir cada sentença, o programa apresentava na tela as três profissões (*médica, técnica*

de enfermagem, faxineira) e o participante deveria clicar em uma das opções. Especificamente, no caso do grupo de participantes com escolaridade irregular, o pesquisador se certificou se conseguiam ler/identificar as três opções. Após clicar em uma resposta, o programa registrava em uma planilha de *Excel* a opção indicada pelo participante e carregava automaticamente o estímulo seguinte.

A significância estatística das variáveis explicativas foi verificada através do teste de qui-quadrado no Programa R. As respostas foram analisadas de duas maneiras: distribuição das respostas por cada variável independente estudada e grau de penalização da variante. Para esta última, foram atribuídos diferentes valores para as respostas dadas pelos participantes aos três perfis de profissões: o valor 1 foi atribuído ao perfil *médica*, o valor 2 foi atribuído ao perfil *auxiliar de enfermagem* e o valor 3 foi atribuído ao perfil *faxineira*. Nesse sentido, o valor 1 foi atribuído às formas entendidas como de maior prestígio e 3 às formas menos prestigiadas a partir da relação entre as variantes produzidas e as profissões as quais os participantes escolhiam. Assim, quanto mais alto o valor atribuído a uma determinada variante, significa que houve mais associações com a profissão de faxineira, e, conseqüentemente com o perfil social de baixa escolaridade e de atividade profissional pouco valorizada. A soma dos valores atribuídos às respostas de todos os participantes revela o grau de estigma ou prestígio das variantes.

Também foi realizada a análise em que a variável dependente foi a escolha da profissional e, nesse caso, as variáveis independentes foram<sup>1</sup>: a) variante (intra-sujeitos/ *within subjects*): glotal ou pós-alveolar; b) número de variantes por sentença (entre-sujeitos/*between subjects*): sentença com 1 ocorrência da variante, sentença com 2 ocorrências da mesma variante; c) e grupo social do participante: UFRJ, Fiocruz, EJLA. A classificação como *within* ou *between subjects* só se aplica a condições dos testes. Esse grupo de fatores é justamente o que mapeia os subjects (ou participantes).

Os objetivos específicos para a realização do teste eram:

- a) verificar se os indivíduos dos diferentes grupos sociais fazem as mesmas avaliações acerca das duas variantes da coda (s). Esperava-se que, em ra-

---

<sup>1</sup> Design experimental intra-sujeitos (*within-subject*) significa que todos os participantes estão expostos aos mesmos tipos de estímulos. Design entre-sujeitos (*between-subjects*) significa que um determinado tipo de estímulo foi exposto a um grupo de participantes e não a outro (DRAGER, 2013, p. 64). Em outras palavras, todos os participantes ouviram as duas variantes de cada variável, mas cada grupo ouviu ou somente uma variante por sentença ou duas variantes por sentença de cada variável sociolinguística avaliada.

zão dos dados de produção obtidos a partir das amostras CENSO 2000, Fiocruz e EJLA, que houvesse uma avaliação negativa – estigmatizada – da fricativa glotal, uma vez que essa variante foi bem menos produzida por falantes com maior escolaridade e pertencentes a diferentes setores da classe média (MELO, 2017);

- b) verificar se a frequência de ocorrência da variante, controlada através da presença de uma ou mais de uma ocorrência da mesma variante no mesmo estímulo, tem efeito na percepção do falante, conforme em Labov et al. (2011).

## RESULTADOS

A fim de verificar se havia associação entre as variáveis testadas, foram realizados testes de qui-quadrado: um p-valor acima de 0,050 indica que há associação entre as variáveis testadas, razão pela qual os resultados obtidos são considerados como tendo significância estatística e, por isso, relevantes no entendimento da percepção das variáveis linguísticas em questão. Serão comentados apenas os resultados que apresentaram relevância estatística no teste do qui-quadrado, quais sejam: associação entre tipo de resposta e ‘grupo social’ e ‘variantes’, de acordo com a metodologia explicitada na seção anterior. Os resultados indicaram que não houve efeito do número de vezes em que a variante aparece em cada sentença, isto é, uma ou duas vezes (p-valor = 0.1868). Apesar de a metodologia utilizada no teste desta pesquisa não ser a mesma daquela utilizada por Labov et al. (2011), os resultados obtidos podem indicar que o grau de estigma da variável testada não tem relação com a frequência de exposição a uma determinada variante ou que a maneira de testar essa variável não tenha conseguido capturar esse efeito.

A Tabela 01 contém os resultados para o grau de estigma atribuído às variantes, considerando o grau de penalização da variante por grupo de participantes.

**Tabela 01:** Resultados: grau de estigma atribuído às variantes do (s) em coda.

	Variantes	
	glotal	pós-alveolar
EJLA	75	78
Fiocruz	113	74
UFRJ	127	57
Total Geral	315	209

De acordo com a Tabela 01, em relação à pontuação geral obtida, verifica-se que houve mais associação das sentenças produzidas com a fricativa glotal ao perfil profissional menos prestigiado socialmente (*faxineira*), tendo sido a pontuação para estas sentenças (315) bem superior às sentenças produzidas com a fricativa pós-alveolar (209). Em relação à distribuição dos índices por grupo social, verifica-se que, quando o estímulo continha a variante glotal, as sentenças não foram bem avaliadas pelos participantes dos grupos UFRJ e Fiocruz (respectivamente, índices de 127 e 113), o que deixa os participantes desses grupos bem próximos em termos de avaliação da variante glotal. Além disso, é possível verificar que não há diferença na avaliação das variantes entre os adolescentes da EJLA, uma vez que os índices atribuídos às variantes glotal e pós-alveolar são muito próximos (respectivamente, 75 e 78). Essa situação coloca os adolescentes do grupo EJLA bem distantes, em termos da avaliação da variante glotal, dos participantes dos outros dois grupos, revelando que as variantes são diferentemente avaliadas dentro da mesma comunidade de fala. O p-valor para as variáveis ‘grupo social’ e ‘variantes’ é de 0.0009345, revelando que a associação entre elas é significativa. Os resultados para a variante glotal mostram também que há uma gradiência na avaliação. Os participantes da UFRJ apresentaram valores de avaliação da glotal mais altos que os da Fiocruz.

A Tabela 02 mostra a distribuição e respectivos percentuais das respostas (associação entre a variante e o perfil de profissão por grupo social) por variante e por grupo de participantes:

**Tabela 02:** Resultados: distribuição das respostas dos participantes para o teste de avaliação do (s) em coda por grupo amostral.

grupo amostral	variantes	Faxineira		Técnica em enfermagem		Médica	
		respostas	%	respostas	%	respostas	%
EJLA	glotal	21	29%	33	46%	18	25%
	pós-alv	19	26%	40	56%	13	18%
Fiocruz	glotal	49	68%	15	21%	8	11%
	pós-alv	23	32%	28	39%	21	29%
UFRJ	glotal	58	81%	11	15%	3	4%
	pós-alv	13	18%	31	43%	28	39%

De acordo com a tabela anterior, comprova-se que as avaliações realizadas em relação às duas variantes pelos participantes dos grupos UFRJ e da Fiocruz se afastam bastante da avaliação feita pelos adolescentes da EJLA. A partir dos percentuais obtidos, verifica-se que, entre os participantes dos grupos UFRJ e Fiocruz, há uma polarização clara das avaliações. Em outras palavras, observa-se que, para os participantes dos grupos UFRJ e Fiocruz, houve uma relação muito forte entre as sentenças realizadas com a variante glotal e o perfil de menor prestígio social (faxineira), o que caracteriza a avaliação negativa que esses participantes fazem do uso da variante glotal.

Por outro lado, entre os participantes da EJLA, não se verifica a mesma polarização, podendo ser percebida, ainda, uma concentração maior para associação entre as sentenças realizadas com a variante glotal e o perfil de técnica em enfermagem, o qual é assumido como um valor central na escala de valor social. Como não há polarização na associação entre a variante glotal e os perfis mais extremos, além de maior concentração no perfil técnica em enfermagem, entende-se que, para os participantes da EJLA, não há uma avaliação negativa ou positiva da variante glotal, diferentemente do que acontece com os outros dois grupos.

A relevância de associação entre as variáveis testadas – ‘grupo social’ e ‘variante glotal’ – é revelada por meio do p-valor encontrado para associação entre elas:  $8.195e-09^2$ . Da mesma forma, o p-valor para a associação de cada grupo social com as duas variantes confirma a análise realizada. Entre os participantes dos grupos UFRJ (p-valor de  $2.294e-13$ ) e Fiocruz (p-valor de  $6.956e-05$ ), há diferentes avaliações das variantes glotal e pós-alveolar. Já entre os participantes do grupo EJLA, não há distinção de avaliação entre as duas variantes (p-valor de 0.4544). Esses resultados refletem os apresentados na Tabela 01.

Se é possível detectar que a variante glotal é a realização menos prestigiada para a variação do (s) em coda entre os participantes dos grupos UFRJ e Fiocruz, não é possível falar em uma variante de prestígio para essa mesma variável. A partir da associação entre os perfis e as variantes testadas, percebe-se que, sobre a variante pós-alveolar, parece não pesar uma avaliação social de prestígio, isto é, os participantes dos três grupos não reconhecem a fricativa pós-alveolar como sendo a variante prestigiada. É possível, então,

---

<sup>2</sup> A notação e-09 no p-valor significa que o resultado encontrado contém, como neste caso, nove casas decimais com zero depois da vírgula e a partir daí os demais números indicados. Essa notação é utilizada para simplificar a apresentação do resultado que seria 0.000000008195, um resultado importante que indica um p-valor muito baixo e, portanto, a forte significância do resultado.

que a variante pós-alveolar funcione como característica regional da variedade falada na cidade do Rio de Janeiro.

## CONSIDERAÇÕES FINAIS

Diferentes estudos sobre a coda (s) em diversas variedades do PB, incluindo a comunidade de fala da cidade do Rio de Janeiro, apontam para um estigma associado à fricativa glotal. Pretendia-se, por meio da aplicação do teste de percepção com estímulos com duas variantes da variável coda (s), atestar este estigma, bem como verificar se essa avaliação é compartilhada por todos os falantes da comunidade de fala do Rio de Janeiro. Por um lado, os resultados confirmaram a avaliação negativa da variante glotal. No entanto, por outro lado, os resultados mostraram que essa avaliação não é compartilhada por todos os falantes da comunidade de fala, uma vez que o estigma da variante glotal foi confirmado entre os participantes mais escolarizados (grupos UFRJ e Fiocruz), mas não entre os adolescentes excluídos socialmente (grupo EJLA). Para este grupo, os resultados indicaram que não houve diferença de avaliação em relação às duas variantes da variável observada.

Uma vez que os participantes dos grupos EJLA e Fiocruz pertencem ao mesmo grupo social (moradores de favelas), é possível que os diferentes graus de inserção social desses indivíduos tenham impacto no comportamento diferenciado em relação à percepção das duas variantes mencionadas. Em razão da situação de exclusão social que vivenciam, os indivíduos do grupo EJLA não se identificam com valores sociais detectados em outros setores da sociedade, o que faz com que, conseqüentemente, não adotem e nem compartilhem tais valores. Por outro lado, a escolaridade regular – todos cursavam o Ensino Médio – e o acesso, por parte dos indivíduos do grupo da Fiocruz, a curso de formação em uma importante instituição de pesquisa, pode levar à adoção desses valores até como uma atitude de diferenciação/distanciamento de outros indivíduos com a mesma origem social, atitude esta que pode ser também observada na taxa geral, nas amostras de fala, de realização da variante glotal para falantes deste grupo, 5%, semelhante à observada para os falantes da Censo 2000, 6%, e distanciada da do grupo EJLA, 30% (MELO, 2012).

Esse quadro mostra que o requisito de compartilhamento de padrões de avaliação no conceito de comunidade de fala de Labov (1972) é limitador e torna mais complexo o enquadramento dos falantes em função de seu comportamento quando observados em relação a mais de uma variável. De acordo com a definição de Labov, considerando os resultados para a avaliação das variantes

da coda (s), os grupos da Fiocruz e da UFRJ podem ser enquadrados como pertencentes à mesma comunidade de fala, ao passo que os falantes da EJLA pertenceriam a outra comunidade de fala. Em suma, é possível que as diferentes experiências sociais dos indivíduos de uma sociedade estratificada socialmente e com diferenças profundas entre alguns segmentos levem ao desenvolvimento de padrões de avaliação diferentes. Portanto, se são consideradas outras variáveis linguísticas, significa que os mesmos grupos sociais se situam igualmente em relação ao compartilhamento de padrões de avaliação social? Se não, tomando por base o conceito de Labov, de quantas comunidades de fala um mesmo falante faz parte? O estudo de Melo (2017) mostrou que, com relação à avaliação de variantes da coda (r), os falantes da UFRJ se diferenciam dos dois grupos de adolescentes (EJLA e Fiocruz) na atribuição de prestígio à realização da coda em interior de palavra, como em *parceiro*. Outra questão que se coloca é o que levaria o falante a apresentar tendências diferentes para diferentes variáveis, como no caso do que ocorre com os participantes da Fiocruz que ora se comportam como os participantes da UFRJ, em relação à avaliação da variante glotal da coda (s) e ora se aproximam dos participantes da EJLA, com relação à avaliação da presença da coda (r). Ainda, conforme observado na Tabela 01, pode haver diferenças gradientes de avaliação entre grupos sociais. Observa-se, então, que a dinâmica sociolinguística de uma determinada comunidade de fala é mais complexa do que a situação reportada em Labov (1972). Conforme já mostrado em outros estudos (SANTA ANA; PARODI, 1998; ZHANG, 2005), propõe-se que o conceito de comunidade de fala integre a diversidade de padrões de avaliação social que pode emergir de forma a capturar, mais amplamente, as diferentes relações que os indivíduos estabelecem em função de seu pertencimento de classe, do grau de inserção social, das suas redes sociais e comunidades de prática das quais participam.

Além disso, o comportamento de cada grupo em relação à percepção da variante velar/glotal parece espelhar os resultados obtidos em Melo (2012, 2017) para dados de produção. Foi observado ainda que os falantes excluídos socialmente da classe baixa (Amostra EJLA) diferem tanto dos falantes da classe média (Amostra Censo 2000) quanto dos falantes da classe baixa com maior inserção social (Amostra Fiocruz) no que se refere à representação detalhada das palavras com a coda (s). Para algumas palavras, a variante glotal é o exemplar dominante entre os falantes da Amostra EJLA, enquanto a variante pós-alveolar é a variante dominante para os outros dois grupos. A mudança em progresso afeta palavras mais frequentes somente entre os falantes da Amostra EJLA. Este grupo é o único que não mostra uma avaliação negativa para a variante velar/glotal.

Também defendemos que o conceito de sistema adaptativo complexo (BECKNER et al., 2009) é mais apropriado para acomodar os resultados obtidos neste estudo. Na abordagem da linguagem como um sistema adaptativo complexo, assume-se que a língua não se desenvolve de maneira uniforme e ordenada. Ao contrário, é dependente das interações específicas dos indivíduos (MUFWENE, 2008, p. 62, 2013, p. 207-208), que estão encaixadas na estrutura da sociedade, uma vez que a experiência do indivíduo é também social e histórica.

## REFERÊNCIAS

AULER, Mônica. A difusão lexical num fenômeno de aspiração em português. In *Revista de Estudos da Linguagem*. Belo Horizonte, Faculdade de Letras da UFMG, 1992.

BECKNER, C.; BLYTHE, R.; BYBEE, J.; CHRISTIANSEN, M. H.; CROFT, W.; ELLIS, N. C.; HOLLAND, J.; KE, J.; LARSEN-FREEMAN, D.; SCHOENEMANN, T. Language is a complex adaptive system: position paper. *Language Learning*, v. 59, s. 1, p. 1-26, 2009.

BRESCANCINI, C. R. A fricativa em posição de coda no PB. In: Jânia M. Ramos. (Org.). *Estudos sociolinguísticos: os quatro vértices do GT da ANPOLL*. Belo Horizonte-MG: Editora da FALE/UFMG, p. 06-20, 2006.

BYBEE, Joan. *Phonology and language use*. Cambridge: Cambridge University Press, 2001.

\_\_\_\_\_. *Language, usage, and cognition*. Cambridge: Cambridge University Press, 2010.

CAMPBELL-KIBLER, Kathryn. The effect of speaker information on attitudes toward (ING). In *Journal of Language and Social Psychology*, vol. 29(2), 214–223, 2010.

CALLOU, Dinah, MARQUES, M. D. *O -s implosivo na linguagem do Rio de Janeiro*. Rio de Janeiro: Littera, 1975.

CARVALHO, Rosana Siqueira de. *Variação do /S/ pós-vocálico na fala de Belém*. Dissertação (Mestrado) – Programa de Pós-Graduação em Linguística, Universidade Federal do Pará, 2000.

CANOVAS, Maria Irene Francisco. *Variação fônica de /s/ pós-vocálico e de /v, z, / cabeças de sílaba na fala da cidade de Salvador*. Dissertação (Mestrado) – Universidade Federal da Bahia. Salvador, 1991.

CLOPPER, Cynthia G, CONREY, Bryanna, PISONI, David. B. Effects of talker gender on dialect categorization. *Journal of Language and Social Psychology*. p. 182-206, 2005.

CONNINNE, C. M.; RANBOM, L. J.; PATTERSON, D. J. Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics*, vol. 70, n. 3, 2008, p. 403–411.

GRYNER, Helena e MACEDO, A. V. T. A pronúncia do –S pós-vocálico na região de Cordeiro – RJ. In: Maria Cecília Mollica; Mário Eduardo Martellota. (Org.). *Análises lingüísticas: a contribuição de Alzira Macedo*. Rio de Janeiro: Serviço de Publicações - FL/UFRJ, p. 26-51, 2000.

GUY, G. As comunidades de fala: fronteiras internas e externas. In: *Abralin*, 2001. Disponível em: <[http://sw.npd.ufc.br/abralin/anais\\_con2int\\_conf02.pdf](http://sw.npd.ufc.br/abralin/anais_con2int_conf02.pdf)>. Acesso em: 10 de setembro de 2019.

HAY, Jennifer, WARREN, Paul; DRAGER, Katie. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, 34(4), p.458-484. 2006.

LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.

\_\_\_\_\_. *The Social Stratification of English in New York City*. New York: Cambridge University Press, 2006.

\_\_\_\_\_. *Padrões Sociolinguísticos*. Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso (tradução). São Paulo: Parábola Editorial, 2008.

LABOV et al. *Journal of Sociolinguistics*. Blackwell Publishing: p. 431-463, 2011.

LAMBERT, W. E et al. Evaluational reactions to spoken languages. In *Journal of Abnormal and Social Psychology*, vol. 60(1), p. 44-51, 1960.

MELO, M. A. S. L. de. *Desenvolvendo novos padrões na comunidade de fala: um estudo sobre a fricativa em coda na comunidade de fala do Rio de Janeiro*. Rio de Janeiro, 2012 Dissertação (Mestrado) UFRJ, Faculdade de Letras, 2012.

MELO, M. A. S. L. de. *Direcionalidade da mudança sonora: o papel do item lexical e da avaliação social*. Rio de Janeiro. Tese (Doutorado em Linguística) – UFRJ, Faculdade de Letras, Rio de Janeiro, 2017.

MOTA, Jacyra Andrade. *O –S em coda silábica na norma culta de Salvador*. Rio de Janeiro: Universidade Federal do Rio de Janeiro, Faculdade de Letras. Tese de Doutorado em Língua Portuguesa, 2002.

MUFWENE, S. *Language Evolution: Contact Competition, and Change*. London: Continuum, 2008.

MUFWENE, S. The Emergence of Complexity in Language: An Evolutionary Perspective. In MASSIP-BONET, A.; BASTARDAS-BOADA, (Eds.) *Complexity Perspectives on Language, Communication and Society*. Berlin: Springer, 2013.

OUSHIRO, Livia. *Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*. Tese (Doutorado em Letras) – USP, FFLCH, São Paulo, 2015.

PIERREHUMBERT, J. B. Probabilistic Phonology: Discrimination and Robustness. In R. Bod, J. Hay and S. Jannedy (eds.) *Probability Theory in Linguistics*. The MIT Press, Cambridge MA, 2003, p.177-228.

RONCARATI, C. N. S. Enfraquecimento das fricativas sonoras. In ARAGÃO, M. do S.; BARROS, K. S. M. (Org.). *Linguística*. Ceará: Universidade Federal do Ceará, v. 1, p. 5-6, 1999.

SANTA ANA, O.; PARODI, C. 1998. Modelling the speech community: Configurations and variable types in the Mexican Spanish setting. In *Language in Society* 27(1), p. 23-51, 1998.

SCHERRE, Maria Marta Pereira e MACEDO, A. V. T. Restrições fonético-fonológicas e lexicais: o -S pós-vocálico no Rio de Janeiro. In: Maria Cecília Mollica; Mário Eduardo Martelotta (org.). *Análises linguísticas: a contribuição de Alzira Macedo*. Rio de Janeiro: Serviço de Publicações - FL/UFRJ, p. 52-64, 2000.

SQUIRES, L. M. *Sociolinguistic priming and the perception of agreement variation: Testing predictions of exemplar-theoretic grammar*. Tese (Doutorado), University of Michigan, 2011.

STAUM CASASANTO, Laura. *What do Listeners Know about Sociolinguistic Variation?* University of Pennsylvania Working Papers in Linguistics: Vol. 15: Iss. 2, Article 6, 2010.

VOTRE, Sebastião. Relevância da variável escolaridade. In MOLLICA, M. C; BRAGA, M. L. *Introdução à Sociolinguística: o tratamento da variação*. São Paulo: Contexto, 2010.

ZHANG, Qying. A Chinese yuppie in Beijing: Phonological variation and the construction of a new professional identity. *Language in Society* 34, p.431-466, 2005.

## ANEXO 1

### Lista de sentenças para o teste de avaliação de itens com o (s) em coda

1 item	2 itens
<p>1. Ele <b>às vezes</b> fica nervoso com facilidade.</p> <p>2. A menina queria <b>mesmo</b> sair de casa.</p> <p>3. Sempre <b>depois</b> do vento, vem a chuva.</p> <p>4. Na aula de amanhã, <b>nós</b> vamos fazer prova.</p> <p>5. O ônibus dava <b>várias</b> voltas sem necessidade.</p> <p>6. Ela precisava de dinheiro, <b>mas</b> não tinha a quem pedir.</p> <p>7. <b>Meus</b> vizinhos sempre reclamam do barulho.</p> <p>8. José colocou a vassoura <b>atrás</b> da porta.</p> <p>9. Somente a fé em <b>Deus</b> move montanhas.</p> <p>10. Maria carregou os livros <b>desde</b> lá de baixo.</p> <p>11. O réu aguardava a sentença do <b>juiz</b> na sala de audiência.</p> <p>12. Quanto <b>menos</b> doce comer, mais rápido fica bom.</p>	<p>1. <b>Às vezes</b> eu tomo <b>meus</b> remédios sem receita.</p> <p>2. <b>Às vezes</b> o jornal dá <b>várias</b> notícias sem sentido.</p> <p>3. <b>Depois</b> do show, <b>nós</b> vamos à praia.</p> <p>4. O <b>juiz</b> não queria, <b>mas</b> decidiu contra o réu.</p> <p>5. <b>Nós</b> não vemos Saulo <b>desde</b> o ano passado.</p> <p>6. <b>Depois</b> de muito tempo, o <b>juiz</b> disse a sentença.</p> <p>7. <b>Meus</b> livros surgiram <b>atrás</b> do armário.</p> <p>8. O <b>mesmo</b> carro deu <b>várias</b> voltas até parar.</p> <p>9. <b>Desde</b> ontem, João não teve <b>mais</b> vontade de sair do quarto.</p> <p>10. <b>Deus</b> nunca dá <b>menos</b> do que você deseja.</p> <p>11. <b>Deus</b> me livre de ter <b>menos</b> dinheiro.</p> <p>12. O bom jogador corre <b>atrás</b> da bola <b>mesmo</b> quando está cansado.</p>

# PERCEPTION OF THE VARIATION OF THE CODA (S) IN THE SPEECH COMMUNITY OF RIO DE JANEIRO ACCESSING THE SOCIAL MEANING OF THE BACK-FRICATIVE VARIANT

*Marcelo Alexandre Silva Lopes de Melo (UFRJ)*  
*Christina Abreu Gomes (UFRJ/CNPq)*

This paper presents the results of a perception test that tried to access the social meaning of the variants of the variable coda (s) by speakers from two distinct social groups in the speech community of Rio de Janeiro (Melo, 2017). The tested variants were [ʃ/ʒ] and [h/ɦ]. In this variety of Brazilian Portuguese, variants of coda include alveolar, post-alveolar, back fricatives (velar and glottal) and the absence of coda, as in me[z]mo, me[ʒ]mo, me[ɦ]mo, meØmo. Previous studies with production data (Callou, Leite, Moraes, 1996; Scherre, Macedo, 2000; Melo, 2012) have shown that post-alveolar variants are the most frequent among all variants and, specially, according to Callou and Brandão (2009), its prevalence, in the speech community, is the result of the implementation of a change. In Melo (2012, 2017), with data from middle class and low-class speakers, living in slums, measured by level of schooling and socioeconomic conditions, it was observed an ongoing change towards the velar/glottal variant among socially excluded low class speakers. In both samples, the velar/glottal variant tends to be used when the coda is followed by a voiced consonant, in word final position, in non-monosyllabic words, in unstressed syllables, with no morphological status, in informal styles and it is subject to lexical conditioning. The overall rate of back fricatives is higher among socially excluded low class speakers without

regular schooling (30%) than middle class speakers (5%) and socially included low class speakers (6%).

This study adopts the view of linguistic variation as in the exemplar-based models (Bybee, 2001, 2010; Pierrehumbert, 2003; Drager; Kitley, 2016), according to which indexed linguistic variation is represented in the lexicon. Since the phonetic detail is stored as part of the wordform, the representation is updated through the speaker's experience with language in many different social, discursive and interactional contexts. In Melo (2012, 2017), it was assumed that the different rates of back fricatives for specific words indicates which variant is dominant in relation to the others in speaker's word representations (Connine; Ranbom; Patterson, 2008). It was shown that, among socially excluded low class speakers, back fricative prevails in some words, such as *mesmo* same, *desde* since, *vários* many (pl.) and *nós* we, whilst, among middle class speakers and socially included low class speakers, these words were mostly produced with the post-alveolar variant. Then, since the organization of words that contain coda (s) is different, at least for some words, among the three groups, it is important to investigate in what extent they share the same patterns of evaluation of these two variants. Different studies, based on production data, provided evidence that the back-fricative variant (velar / glottal) is stigmatized in several varieties of Brazilian Portuguese (Auler, 1992; Carvalho, 2000; Gryner, Macedo, 2000).

Perception of variants was assessed using a matched-guise method. The perception test contained 24 sentences constructed according to the number of words with the relevant coda in the sentence: 12 sentences with only one word (*Sempre depois do vento, vem a chuva*/When the wind stops blowing, there comes the rain), and 12 sentences with two words with the relevant coda (*Depois de muito tempo, o juiz disse a sentença*/After a long while, the judge gave the sentence). The sentences were controlled by size, coda's following context, and distance between words in the "two words with coda (s)" condition. All sentences were recorded by the same woman, from the same speech community of the participants, aged between 25-30 years old, and with university level of education, in two versions: one with the post-alveolar and the other with the velar/glottal. Stimuli were split in two lists of 12 sentences, one containing sentences with only one word with coda (s) and the other with two words. The number of words with coda (s) was an across-subject design, since participants heard stimuli from one of either lists. In each list, half of the sentences contained the post-alveolar variant and the other half, the velar/glottal variant, which means that no participant heard the same word with two different variants. Each participant heard 12 sentences for the relevant coda and 06 fillers.

The task consisted of the indication of the professional who said the stimuli among three options of occupation: physician (professional with a university degree required), nurse (professional/skilled worker, but with no university degree required) or cleaner (unskilled). These professionals were selected for testing the correlation between variant and social meaning because of their different social values aligned in a scale of different degrees of prestige in Brazilian society, according to the type of specialization and level of education required to practice them. The participants were instructed to listen each sentence carefully and then select one of the three professionals shown on the screen they thought that have said it. The familiarization with the task consisted in the presentation of a text on the screen, read by the researcher, in which it was told that the participant was at a public hospital, waiting for an appointment. While he/she was waiting, the participant heard a sentence, and had to decide whether this sentence was produced by a physician, by a nurse or by a cleaner (unskilled). The test was applied individually in the institutions where the participants were contacted using the software TP (Test/Training of Perception), version 3.1. The stimuli were presented randomly and the answers were registered on/in an Excel sheet for further analysis.

The test was applied to 36 individuals, distributed in three sets according to level of schooling, which also indicates different degrees of social insertion: a) UFRJ group, constituted by university students (from the Federal University of Rio de Janeiro), mainly with a middle and low-middle class profile; b) Fiocruz group, constituted by high school slum-living adolescents that participate in an extra-class activity offered by an important public research institution; c) EJLA group, constituted by socially excluded slum-living adolescents with no access to formal education, living in a reformatory.

The responses were analyzed in two ways: attributing a value to each profession in a scale from 1 to 3, in which 1 – doctor, 2 – nurse, 3 – cleaner per participant' social group, and through the distribution of the answers according all the independent variables. In this case, the dependent variable was the chose professional, and the independent variables were the variant (post-alveolar x velar/glottal), the number of words in the sentence with the relevant coda (1 x 2) and participant' social group (UFRJ x Fiocruz x EJLA).

The results for the participants' choice according the evaluation scale from 1 to 3 (the sum of the points obtained in each answer according to the chose professional) showed that the difference among the three groups and the two variants are statistically significant (chi-square test – p-value=0.0009345). The glottal va-

riant presented score (315) well above the post-alveolar (209), which means more association of this variant to cleaner. It was observed a gradual increase of the scores for the glottal variant from the socially excluded participants to the university participants, respectively, EJLA – 75, Fiocruz – 113, UFRJ – 127, whilst the scores for the post-alveolar don't differ deeply among the groups, respectively, between the lowest and the higher values.. These results indicate that for EJLA participants there is no difference of evaluation between the two variants, and that the social evaluation of the glottal differs gradually among the three groups.

The distribution of the responses per variant per participant's group complements the results obtained by considering the degree of penalization of the variant according to the chose professional. Distribution of responses of EJLA participants showed that there is no correlation between any of the two variant to a profession (p-value=0.4454), whilst there is a polarization between the two opposite professionals – physician and cleaner, among Fiocruz and UFRJ participants, respectively, Fiocruz: 11%, 68% (p-value=6.956e-05) and UFRJ: 4%, 81% (p-value= 2.294e-13) for the glotal fricative. For these two groups, the distribution of responses for the post-alveolar variant showed no bias related to the degree of prestige of the professional (Fiocruz: physician: 29%, nurse: 39%, cleaner: 32%; UFRJ: physician: 39%, nurse: 43%, cleaner: 18%). These results indicate that the post-alveolar is not associated to a value of prestige, and that it is probably perceived as a feature that characterizes the variety spoken in Rio de Janeiro.

Taken together, the results obtained from the different analysis of the perception of the two variants of coda (s) revealed that the three observed social groups don't share the same patterns of social evaluation, that there is a bias to the stigmatization of the glottal variant but no bias to evaluate the post-alveolar as a prestige form. In addition, the behavior of each group in the perception of the glottal variant seems to mirror the results obtained in Melo (2012, 2017) for production data. It was observed that socially excluded low class speakers differ from middle class speakers and socially included low class speakers in terms of the detailed representation of words with coda (s). For some words, the glottal variant is the dominant exemplar among EJLA speakers, whilst the post-alveolar is the dominant variant for the other two EJLA groups of speakers. The ongoing change affects high frequent words only for the EJLA group. This group is the one that doesn't show a negative evaluation for the velar/glottal variant.

The perception of coda (s) study also revealed that the sociolinguistic dynamics of speakers of a given speech community is complex, since there are not only different evaluations of the coda variants among the different groups of

participants, but also that these differences are gradual. The results, moreover, pointed out that evaluation patterns are closely related to the degree of access of the individuals of the different groups to the institutions that contributes to the maintenance and implementation of the prestige patterns.

We also argue that the concept of dynamic adaptive system (Beckner et al., 2009) is more appropriate to accommodate the results obtained by this study. In approaching language as an adaptive and dynamic system, it is assumed that language does not evolve in a uniform and orderly manner, depending on the specificities of individuals' interactions. (Mufwene, 2008: 62; 2017:207-208). However, it is clear that linguistic experience is embedded in the social structure, since individual experience is also social and historical.

## REFERENCES

AULER, Mônica. A difusão lexical num fenômeno de aspiração em português. In *Revista de Estudos da Linguagem*. Belo Horizonte, Faculdade de Letras da UFMG, 1992.

BYBEE, Joan. *Phonology and language use*. Cambridge: Cambridge University Press, 2001.

\_\_\_\_\_. *Language, usage, and cognition*. Cambridge: Cambridge University Press, 2010.

BOERSMA, Paul; WEENINK, David. Praat: doing phonetics by computer, 2018. Disponível em: <<http://www.fon.hum.uva.nl/praat/>>. Acesso em: 26 jan. 2019.

CALLOU, D. M.; MORAES, J.; LEITE, Y. “Apagamento do /R/ final no dialeto carioca: um estudo em tempo aparente e em tempo real”. *D.E.L.T.A.* vol. 14, p. 61-72, 1996.

CALLOU, D. M. I.; BRANDÃO, S. F. “Sobre o /S/ em coda silábica no Rio de Janeiro: falas culta e popular”. In: SALGADO, Ana Claudia Peters; SAVEDRA, Mônica M. Guimarães. (Org.). *Sociolinguística no Brasil: uma contribuição dos estudos sobre línguas em/de contato: homenagem ao Prof. Jürgen Heye*. Rio de Janeiro: 7 Letras, p. 27-34, 2009.

DRAGER, K.; KIRTLEY, M. J. Awareness, Salience, and Stereotypes in Exemplar-Based Models of Speech Production and Perception. In: BABEL, A. (Org.). *Awareness and Control in Sociolinguistic Research*. Cambridge: Cambridge Press, 2016. p. 1-24.

HELLWIG, B.; GEERTS, J. ELAN – Linguistic Annotator. Versão 5.4. Disponível em: <<http://www.mpi.nl/corpus/manuals/manual-elan.pdf>>, 2018. Acesso em: 26 jan. 2019.

GOLDMAN, Jean-Philippe. *EasyAlign*: an automatic phonetic alignment tool under Praat. In: Proceedings of InterSpeech, 2011. Disponível em: <<https://archive-ouverte.unige.ch/unige:18188>>. Acesso em: 30 set. 2016.

KENDALL, Tyler; THOMAS, Erik R. Package ‘vowels’, 2015. Pacote para a plataforma R. Disponível em: <<http://blogs.uoregon.edu/vowels/2012/11/08/vowels-r-1-2/>>. Acesso em: 30 set. 2016.

LOBANOV, Boris M. Classification of Russian vowels spoken by different speakers. *Journal of Acoustic Society of America*, vol. 49, n. 2, 606-608, 1971.

OUSHIRO, Livia. *silac: Transcritor fonológico do português*. 2018. Versão online (v0.5.1). Disponível em: <[oushiro.shinyapps.io/silac](http://oushiro.shinyapps.io/silac)>. Acesso em: 18 jan. 2019.

RIEBOLD, John. *Vowel analyzer*. Ms., 2013. (script do Praat). Disponível em: <<https://raw.githubusercontent.com/jmriebold/Praat-Tools/master/Vowel-Analyzer.praat>>. Acesso em: 30 set. 2016.

CARVALHO, Rosana Siqueira de. *Variação do /S/ pós-vocálico na fala de Belém*. Dissertação (Mestrado) – Programa de Pós-Graduação em Linguística, Universidade Federal do Pará, 2000.

CONNINNE, C. M.; RANBOM, L. J.; PATTERSON, D. J. Processing variant forms in spoken word recognition: The role of variant frequency. *Perception & Psychophysics*, vol. 70, n. 3, 2008, p. 403–411.

GRYNER, Helena e MACEDO, A. V. T. A pronúncia do –S pós-vocálico na região de Cordeiro – RJ. In: Maria Cecília Mollica; Mário Eduardo Martellotta. (Org.). *Análises lingüísticas: a contribuição de Alzira Macedo*. Rio de Janeiro: Serviço de Publicações - FL/UFRJ, p. 26-51, 2000.

MELO, M. A. S. L. de. *Desenvolvendo novos padrões na comunidade de fala: um estudo sobre a fricativa em coda na comunidade de fala do Rio de Janeiro*. Rio de Janeiro, 2012 Dissertação (Mestrado) UFRJ, Faculdade de Letras, 2012.

MELO, M. A. S. L. de. *Direcionalidade da mudança sonora: o papel do item lexical e da avaliação social*. Rio de Janeiro. Tese (Doutorado em Linguística) – UFRJ, Faculdade de Letras, Rio de Janeiro, 2017.

MUFWENE, S. *Language Evolution: Contact Competition, and Change*. London: Continuum, 2008.

PIERREHUMBERT, J. B. Probabilistic Phonology: Discrimination and Robustness. In R. Bod, J. Hay and S. Jannedy (eds.) *Probability Theory in Linguistics*. The MIT Press, Cambridge MA, 2003, p.177-228.

SCHERRE, Maria Marta Pereira e MACEDO, A. V. T. Restrições fonético-fonológicas e lexicais: o -S pós-vocálico no Rio de Janeiro. In: Maria Cecília Mollica; Mário Eduardo Martelotta (org.). *Análises lingüísticas: a contribuição de Alzira Macedo*. Rio de Janeiro: Serviço de Publicações - FL/UFRJ, p. 52-64, 2000.



## QUESTÕES E MÉTODOS VOGAIS MÉDIAS PRETÔNICAS NA FALA DE MIGRANTES NORDESTINOS EM SITUAÇÃO DE CONTATO DIALETAL

*Livia Oushiro (Unicamp)*

### RESUMO

Este capítulo tem o duplo objetivo de analisar os padrões de variação na altura das vogais médias pretônicas /e/ e /o/ (p.ex., *relógio* e *romã*) na fala de 32 migrantes paraibanos e alagoanos residentes na cidade de São Paulo, e discutir métodos para o tratamento de dados sociofonéticos. Tendo em vista que, em situação de contato entre dialetos, mudanças na fala dos indivíduos podem ocorrer de modo gradual, as vogais são analisadas como variáveis contínuas (medidas de F1 em Hz), em contraste com os padrões de 7 paulistanos. Os dados foram automaticamente extraídos, medidos e codificados por meio de *scripts* no Praat (BOER-SMA; WEENINK, 2018) e R (R CORE TEAM, 2018). Em análises de regressão linear de efeitos mistos, os resultados mostram que, de 6 variáveis predictoras sociais, apenas a Idade de Migração se correlaciona significativamente com a altura das vogais /e/ e /o/ – quanto mais cedo migrou, mais altas as vogais pretônicas. Ao discutir os resultados das análises à luz dos métodos empregados, avalia-se que o campo de estudos variacionistas no Brasil tem limitado seu escopo de questões devido à dependência de um único modelo estatístico de análise, o que requer uma mudança de postura em relação aos dados sociolinguísticos.

## INTRODUÇÃO<sup>1</sup>

Desde os primeiros trabalhos da Sociolinguística Variacionista (LABOV, 2008 [1972]), as variáveis de natureza fonética e fonológica têm tido destaque. No contexto da Sociolinguística brasileira, especificamente, já são inúmeros os trabalhos sobre a realização variável de /r/ em coda silábica (p.ex., CALLOU; MORAES; LEITE, 1996; OLIVEIRA, 1999; BRESCANCINI; MONARETTO, 2008), /s/ em coda (p.ex., MACEDO, 2004; BRANDÃO, 2008), palatalização de /t/ e /d/ antes de /i/ (p.ex., BATTISTI; DORNELLES FILHO, 2012; MARTINS, 2008), o alçamento e o abaixamento de vogais médias pretônicas (p.ex., BORTONI; GOMES; MALVAR, 1992; YACOVENCO, 1993; PEREIRA, 1997), entre outras. Para esse tipo de variável, não raro se obtêm centenas de dados a partir de cerca de uma hora de gravação por falante, algo que nem sempre ocorre para variáveis de natureza morfológica ou sintática (OUSHIRO, 2011, p. 22). Esse fato levanta a questão de como trabalhar com grandes quantidades de ocorrências de modo sistemático e mais eficiente.

Além disso, variáveis fonético-fonológicas mormente se caracterizam por sua natureza gradual e contínua em vez de categorias discretas que não se sobrepõem. Contudo, as análises no programa Varbrul (em suas diversas implementações) não permitem capturar tais gradações. Em sua última versão, Goldvarb X (SANKOFF; TAGLIAMONTE; SMITH, 2005), o programa permite apenas a inclusão de variáveis dependentes categóricas e binárias, i.e., a contraposição de duas variantes *x* vs. *y* (p.ex., vogal média baixa vs. vogal média alta), e variáveis independentes igualmente fatoriais (p.ex., faixa etária 1, faixa etária 2 etc.). Nele, não é possível, por exemplo, analisar a altura das vogais médias pretônicas com base em suas medidas de F1 ou analisar uma possível correlação com a idade do falante medida em anos.

Dentro desse cenário, o presente capítulo tem dois objetivos. O primeiro é o de apresentar os métodos empregados na análise da altura variável de vogais médias pretônicas (como em “r/e/lógio” e “r/o/mã”) na fala de 32 migrantes alagoanos e paraibanos residentes na cidade de São Paulo, em comparação com

---

<sup>1</sup> Os resultados desta pesquisa foram frutos de dois financiamentos: a metodologia foi desenvolvida durante o estágio de pós-doutorado realizado no Departamento de Linguística da Universidade Federal do Rio de Janeiro em 2015-2016, sob supervisão de Christina Abreu Gomes, com bolsa do Programa Nacional de Pós-Doutorado (PNPD/CAPES); as análises dos dados de 32 migrantes foram desenvolvidas durante o primeiro ano do Projeto Regular “Processos de acomodação dialetal na fala de nordestinos residentes em São Paulo”, sob os auspícios da FAPESP (Processo 2016/04960-7).

padrões de fala de paulistanos nascidos e criados na capital paulista. Em situações de contato dialetal, é possível imaginar que um migrante nordestino no Sudeste passe a produzir vogais médias relativamente mais altas do que as de seu dialeto de origem, ainda que relativamente mais baixas do que as dos falantes da comunidade anfitriã. De experiências cotidianas, não é raro ouvir casos de migrantes que dizem ter sido admoestados por conterrâneos pelo fato de terem “mudado seu modo de falar”, mas que, para os ouvidos da nova comunidade, ainda têm “um forte sotaque”. A altura das vogais constitui, na realidade, uma variável contínua (definida principalmente pelas medidas do primeiro e do segundo formantes em Hertz), com uma série de graus intermediários entre as vogais prototípicas de cada variedade do Português.

Ao apresentar os resultados de tais análises, o segundo objetivo visa a chamar a atenção para duas limitações recorrentes em estudos variacionistas, para os quais, argumenta-se, os sociolinguistas devem ativamente buscar soluções: (i) o fato de que, não raro, o tempo empregado nas tarefas mecânicas de identificação, extração e codificação dos dados (sobretudo para variáveis fonético-fonológicas) ultrapassa o tempo dedicado às análises em si, verdadeiro foco do pesquisador; e (ii) a subordinação das questões de pesquisa a um conjunto restrito de métodos de análise, relação que deveria ser subvertida: são as questões de pesquisa e as preocupações teóricas que devem guiar a escolha de métodos de análise. Tais impasses só serão superados por meio da valorização e da divulgação de métodos empregados (muitas vezes considerados secundários em um estudo), assim como pela atualização constante do pesquisador quanto aos recursos disponíveis naquilo que Tagliamonte (2012, p. 137) chamou de “caixa de ferramentas da Sociolinguística Variacionista”.

Com esses objetivos em mente, este trabalho se encontra organizado do seguinte modo: na Seção 1, apresentam-se detalhadamente os métodos e ferramentas empregados no estudo da altura variável de vogais médias pretônicas na fala de migrantes, com vistas à sua divulgação; em seguida, relatam-se e discutem-se os resultados da análise sobre as vogais em correlação com fatores sociais<sup>2</sup> (Seções 2 e 3), com especial atenção à pertinência dos métodos; por fim, conclui-se o artigo com as possibilidades de emprego das mesmas ferramentas em estudos sociofonéticos futuros e a importância desses recursos para o desenvolvimento da área.

---

<sup>2</sup> Por concisão, aqui não se apresentam os resultados de correlação com variáveis linguísticas. Contudo, cabe notar que tais variáveis também foram incluídas nos modelos estatísticos.

## MATERIAIS E MÉTODOS

O *corpus* desta análise consiste em 32 gravações de entrevistas sociolinguísticas, cada qual com cerca de uma hora, com 21 migrantes alagoanos e 11 paraibanos residentes na cidade de São Paulo.<sup>3</sup> Ambas as amostras foram estratificadas de acordo com o sexo/gênero do falante (masculino, feminino), três faixas etárias (20-34, 35-59, 60 ou mais anos), e dois níveis de escolaridade (até Ensino Médio, Nível Superior). A partir das gravações, também se depreenderam outras variáveis de interesse para o estudo da fala de migrantes: idade e motivo da migração, tempo de residência em São Paulo, local de origem (rural ou urbano).

Para comparação com a fala paulistana, utilizaram-se os dados de 7 informantes do *corpus* do Projeto SP2010 (MENDES; OUSHIRO, 2012) – três mulheres e quatro homens de perfis sociolinguísticos similares aos dos migrantes. Todas as gravações passaram pelos procedimentos que se descrevem a seguir.

## TRANSCRIÇÃO DE GRAVAÇÕES NO PROGRAMA ELAN

Todas as entrevistas foram, inicialmente, transcritas no programa ELAN (HELLWIG; GEERTS, 2018),<sup>4</sup> de acordo com os procedimentos do Projeto “Processos de acomodação dialetal na fala de nordestinos residentes em São Paulo” (OUSHIRO, 2016). Dentre os recursos do programa, o principal certamente é a possibilidade de criar transcrições alinhadas à gravação, com marcação do tempo inicial e final de cada trecho de fala, o que facilita enormemente a rápida localização de ocorrências de variáveis fonéticas. O programa também permite separar os enunciados de diferentes participantes (como o documentador, o informante e terceiros), além de realizar anotações diversas (como dados contextuais), por meio de trilhas distintas. Por fim, o ELAN permite exportar a transcrição em diversos formatos (.txt, .csv etc.), dentre os quais o mais pertinente para o que segue é o formato .TextGrid do Praat (BOERSMA; WEENINK, 2018). A Figura 1 mostra a interface do programa ELAN: na parte superior, visualizam-se as anotações de uma das trilhas (a fala de “S1”, o participante), com seus

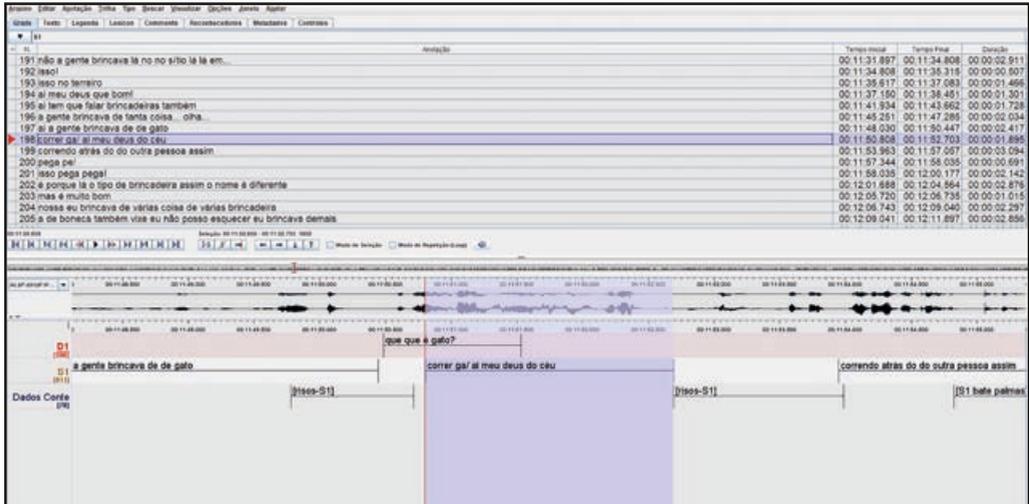
---

<sup>3</sup> Esses dados foram coletados, respectivamente, para a pesquisa de Gomes da Silva (2014), sobre a concordância nominal por parte de migrantes alagoanos em São Paulo, e para o Projeto Casadinho (HORA; NEGRÃO, 2011), uma parceria entre o Programa de Pós-Graduação em Linguística da Universidade Federal da Paraíba e o Programa de Pós-Graduação em Semiótica e Linguística Geral da Universidade de São Paulo, desenvolvida com o apoio do CNPq no âmbito do Edital de Cooperação Nacional n. 16/2008.

<sup>4</sup> Ver tutorial do ELAN em Oushiro (2014).

respectivos tempos iniciais, finais e de duração; na parte central, veem-se os botões de comando para a audição da gravação (como *play*, pausa e diferentes intervalos de recuo ou avanço do áudio), bem como a onda sonora; e, na parte inferior, observam-se as trilhas específicas deste arquivo de transcrição (as trilhas do documentador D1, da informante S1 e de Dados Contextuais).

Figura 1: Exemplo de transcrição no programa ELAN



Fonte: elaboração própria.

Após essa etapa, as transcrições foram exportadas em formato .txt para que a trilha do informante passasse pelo transcritor *silac* (OUSHIRO, 2018).

## DESENVOLVIMENTO DO TRANSCRITOR *SILAC*

O transcritor *silac* (OUSHIRO, 2018) é um *script*, escrito para a plataforma R (R CORE TEAM, 2018), que faz a silabificação, acentuação e transcrição fonológica de textos transcritos em ortografia padrão do Português Brasileiro (PB). Ele é de uso gratuito e está atualmente disponível na forma de aplicativo na Internet em <<https://oushiro.shinyapps.io/silac/>>.<sup>5</sup>

O aplicativo requer que se carregue um arquivo .txt (UTF-8) com o texto que se deseja converter fonologicamente. De modo geral, as regras de transcrição fonológica seguem as convenções do IPA (International Phonetic Alphabet), mas foram estabelecidas certas normas específicas, sobretudo quanto a caracteres especiais:

<sup>5</sup> Último acesso em 26 jan. 2019.

- (i) vogais nasais são precedidas do til: a-'s~i para *assim*, ~e-t~aw para *então*
- (ii) diferenciam-se o “r” forte (como em *arroz*, *rato*, representado por “h”: a-'hos, 'ha-to), o “r” fraco (como em *caro*, *prato*, representado por “r”: 'ka-ro, 'pra-to) e o “r” em coda silábica (como em *quartirão*, representado por “R”: kwaR-tej-'r~aw)
- (iii) os dígrafos “nh” e “lh” são representados por “N” e “L” respectivamente: 'mi-Na para *minha*, 'fa-La para *falha*
- (iv) o fonema /ʒ/ é representado por “Z”: 'Z~e-te para *gente*, 'Zi-ra para *gira*
- (v) o fonema /ʃ/ é representado por “S”: 'Sow para *show*, paj-'S~aw para *paixão*

Um exemplo do resultado da aplicação do *script* se encontra no excerto em (1), que apresenta à esquerda a transcrição ortográfica e à direita a respectiva transcrição fonológica gerada pelo *silac*:

(1)

Transcrição ortográfica	Transcrição fonológica
ai eu nem peguei o meu	'aj 'ew 'n~e pe-'gej o 'mew
vou pegar porque se tocar	'vow pe-'gaR poR-'ke se to-'kaR
Oi	'oj
I. minha filha	'i 'mi-Na 'fi-La
transferir o quê?	tr~as-fe-'riR o 'ke
eu não acredito	'ew 'n~aw a-kre-'di-to
que chatinha hein	ke Sa-'ti-Na e-'~i
ó então fica com o meu celular	'o ~e-'t~aw 'fi-ka k~o o 'mew se-lu-'laR
que acho que até oito horas eu termi- no	ke 'a-So ke a-'te 'oj-to 'o-ras 'ew teR-'mi-no
se tocar você atende pra mim	se to-'kaR vo-'se a-'t~e-de pra 'm~i
Não	'n~aw
pode atender	'po-de a-t~e-'deR
ah eu não não não	a 'ew 'n~aw n~aw 'n~aw
(pede pra ver quem é)	'pe-de pra 'veR 'k~e 'e
o telefone está lá?	o te-le-'fo-ne es-'ta 'la
está com você? fica com você então	es-'ta k~o vo-'se 'fi-ka k~o vo-'se ~e-'t~aw

Cabe ressaltar que, em se tratando de uma transcrição fonológica, o transcritor não discrimina fenômenos variáveis do PB, como o alicamento de vogais postônicas (p.ex. /o/ ~ [u], como em *rato*, *cavalo*), a realização variável do segmento /r/ em coda silábica (p.ex. [maR] para *mar*), a redução do verbo *estar* (p.ex. *ele tá fazendo*), entre muitas outras variáveis. No entanto, é justamente a padronização da transcrição na forma fonológica que facilita a busca de ocorrências para um estudo que se debruce sobre um fenômeno específico; por exemplo, um pesquisador que esteja analisando a pronúncia variável de /r/ em coda silábica (como *tepe*, *retroflexo*, *aspirada*, *apagada* etc.) pode fazer uma busca pelo símbolo “R” na transcrição fonológica – que corresponde, unicamente, a tal segmento nessa posição da sílaba – e checar, junto à gravação, qual realização foi empregada em cada ocorrência. Assim, a transcrição fonológica permite a rápida identificação de ocorrências, principalmente se junto à transcrição alinhada ao arquivo de áudio no programa ELAN.<sup>6</sup>

Para a análise de vogais pretônicas, especificamente, elaborou-se uma sequência adicional de código no *script* para o R, denominada *silacpret*, que, após a transcrição fonológica, identifica e marca as ocorrências dessas vogais entre os símbolos “< >”; assim, palavras como “morar” e “telefone”, por exemplo, são transcritas como “m<o>-raR” e “t<e>-l<e>-fo-ne” respectivamente. A transcrição fonológica no *silacpret* foi importada a uma trilha específica do programa Praat.

## EMPREGO DO *PLUGIN EASYALIGN* E DO *SCRIPTVOWEL ANALYZER*

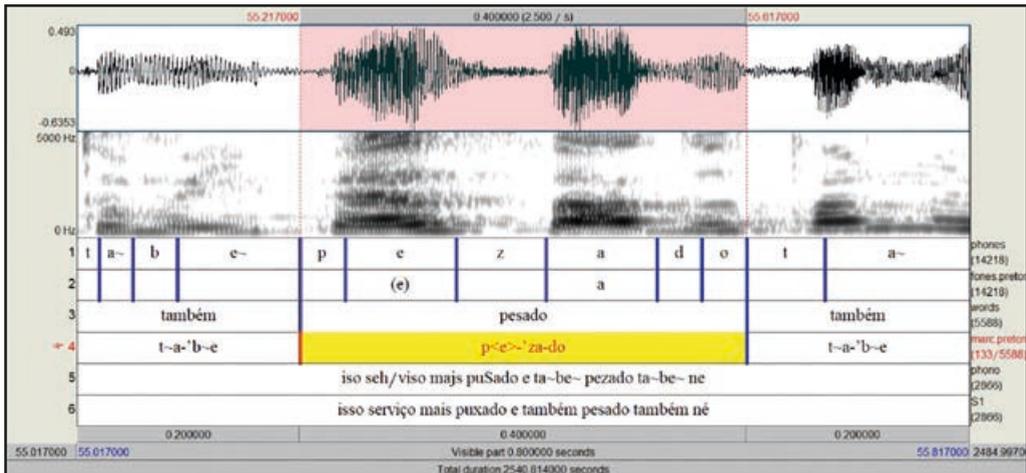
O EasyAlign (GOLDMAN, 2011) é um *plugin* que se instala no programa Praat e realiza, dentre outras tarefas, a segmentação fônica de dados de fala a partir de uma transcrição fonológica, identificando o tempo inicial e final de cada segmento junto à onda sonora.<sup>7</sup> A partir de tal segmentação, o EasyAlign

<sup>6</sup> Há, por outro lado, algumas limitações: (i) palavras estrangeiras normalmente não são transcritas corretamente, uma vez que as convenções ortográficas de outras línguas não seguem aquelas do PB; e (ii) o transcritor não discrimina as vogais média-baixas /e o/ das média-altas /e o/. Para certos itens lexicais, tal diferenciação pode ser feita apenas se se tem informação de sua classificação morfológica e função sintática numa dada sentença -- p.ex., *almoço* como substantivo (/mo/) ou como verbo (/mo/). Ainda que outros tantos casos sejam previsíveis a partir da ortografia (p.ex., para *pastel* /pas-<sup>h</sup>tew/), preferiu-se padronizar a transcrição de vogais médias apenas como “e” e “o” até que se encontre uma solução para todos os casos.

<sup>7</sup> As outras tarefas realizadas pelo EasyAlign são (i) a macrossegmentação de sequências de transcrição, i.e., identifica, em uma cadeia sonora, o ponto inicial e final de sentenças transcritas; e (ii) a fonetização, i.e., a transcrição fonológica a partir de uma transcrição orto-

identifica, ademais, o ponto inicial e final de cada item lexical. A aplicação dessa função do EasyAlign agilizou a localização e segmentação das vogais junto à transcrição e à onda sonora no Praat. A Figura 2 mostra um arquivo de som junto à sua transcrição no programa Praat após a segmentação pelo EasyAlign.

**Figura 2:** Transcrição no Praat após a aplicação do EasyAlign e marcação de vogais pretônicas



Fonte: elaboração própria.

A trilha 6, “S1”, corresponde à fala do participante tal como havia sido exportada do programa ELAN.<sup>8</sup> A trilha 5, “phono”, corresponde à transcrição fonológica do EasyAlign; note-se que o *plugin* segue convenção própria: por exemplo, o “r” em coda é indicado por “h/” em vez de “R”. As trilhas 3 e 4 correspondem à segmentação lexical nas formas ortográficas e fonológica, respectivamente. A trilha 4 corresponde àquela criada por meio da aplicação do *silacpret*, no R, a fim de identificar os segmentos de vogais pretônicas; note-se que a vogal /e/ pretônica de *pezado* está indicada por “<>” nessa trilha. A trilha 1 é o resultado da segmentação do EasyAlign que, dentro de cada sentença, localizou o ponto inicial e final de cada segmento fonológico – e, de especial interesse para a presente análise, os pontos iniciais e finais das vogais pretônicas. Com base

gráfica. A aplicação da função (i) era desnecessária, uma vez que as transcrições já estavam alinhadas por terem sido feitas no programa ELAN (ver Figura 1 acima). A função (ii) realiza tarefa similar ao *script* silacpret acima descrito mas, como não realiza a marcação de vogais pretônicas, deu-se preferência à utilização do *script* de autoria própria.

<sup>8</sup> As trilhas com a fala do documentador, de terceiros e de dados contextuais foram descartadas nesse passo da análise.

nesses dados, foi criada a trilha 2, que separa os dados de vogais pretônicas de interesse para este estudo e as respectivas vogais da sílaba seguinte – no exemplo do item lexical “pesado”, a vogal pretônica /e/, anotada entre parênteses, e a vogal seguinte /a/, esta última para que pudesse ser analisada como variável previsora independente. Além das vogais médias /e/ e /o/, também foram marcadas as demais vogais pretônicas /i, a, u/, para compor todo o quadro vocálico pretônico de cada falante (ver Seção 2).

Após a segmentação de vogais pretônicas, aplicou-se o *script* Vowel Analyzer (RIEBOLD, 2013), no Praat, para extração automática de medições de F1, F2, duração de segmentos e respectivos itens lexicais. O *script* pode ser aplicado não apenas a uma gravação, mas a um conjunto delas, para o qual basta que estejam em uma mesma pasta dentro do computador. A Figura 3 mostra a janela que se abre ao rodar esse *script*; nela, especificam-se: (i) os diretórios em que se encontram os arquivos de gravação e de transcrição no formato .TextGrid (a notação “./” indica que é o mesmo diretório em que está o arquivo do *script*); (ii) o nome do arquivo de resultados a ser gerado (neste caso, “pretonicas-F.txt”); (iii) se se deve utilizar um arquivo de palavras-alvo (neste caso, não); (iv) a “Vowel tier”, trilha que contém a marcação de vogais (aqui, a trilha “fones. pretonicas”); (v) a “Word tier”, trilha que contém os itens lexicais (no exemplo, a trilha “words”); (vi) se se deve extrair um “Notes tier” (neste caso, extraíram-se as informações da trilha de transcrição fonológica); (vii) os pontos de medição de formantes (escolheu-se a opção “30%/50%/70%”); e mantiveram-se as demais opções no formato *default*.

Figura 3: Janela do *script* Vowel Analyzer

The screenshot shows a dialog box titled "Run script: Vowel Analyzer" with a close button (X) in the top right corner. The dialog is organized into several sections:

- Paths:**
  - Soundfile directory:
  - Textgrid directory:
  - Results file:
  - Use targets file:
  - Targets file:
- Tiers:**
  - Vowel tier:
  - Use word tier:
  - Word tier:
  - Use notes tier:
  - Notes tier:
- Formant settings:**
  - Measurement points:
  - Maximum formant (Hz):
  - Number of formants:
- Pitch Settings:**
  - Extract pitch:
  - Pitch range (Hz):
- Analyst:**
  - Initials:

At the bottom of the dialog, there are four buttons: "Standards", "Cancel", "Apply", and "OK".

Fonte: elaboração própria.

O *script* foi rodado em dois conjuntos de gravações, separados pelo sexo dos falantes, pois é recomendável especificar o máximo formante diferentemente para cada sexo. Na Figura 3, esse valor foi especificado como “5.500 Hz” para falantes do sexo feminino. Nos presentes dados, o *script* Vowel Analyzer levou cerca de 5 minutos para extrair as medidas de formantes para cada gravação, com cerca de 180 vogais pretônicas e suas respectivas vogais seguintes anotadas



posta. Sobre a planilha gerada pelo Vowel Analyzer, aplicou-se um novo *script* elaborado na plataforma R para a codificação automática das seguintes variáveis independentes previsoras:

Variáveis linguísticas:

- i) Contexto Fônico Precedente
- ii) Contexto Fônico Seguinte
- iii) Vogal da Sílabas Seguinte (/a/, /ε/ etc.): (a) F1 da vogal da sílabas seguinte (em Hz); (b) F2 da vogal da sílabas seguinte (em Hz)
- iv) Vogal Tônica (/a/, /e/, /i/ etc.)
- v) Distância da Vogal Tônica (em número de sílabas)
- vi) Estrutura da Sílabas Pretônica (CV, CVC, CCV etc.)

Variáveis sociais:

- vii) Amostra (ALSP, PBSP, SP2010)
- viii) Informante
- ix) Sexo (F, M)
- x) Nível de Escolaridade (até Ensino Médio, Nível Superior)
- xi) Faixa etária (20-34, 35-49, 60+) e Idade (em anos)
- xii) Idade de Chegada a São Paulo (em anos)
- xiii) Tempo de Residência em São Paulo (em anos)
- xiv) Motivo da migração (estudo, família, qualidade de vida, trabalho)

São seis variáveis linguísticas frequentemente analisadas em estudos sobre vogais médias pretônicas e oito variáveis sociais de interesse a este estudo. As variáveis Contexto Fônico Precedente, Contexto Fônico Seguinte e Estrutura da Sílabas Pretônica foram codificadas a partir dos dados das colunas “Vowel” e “Transc.Fon”: uma vez identificada a vogal pretônica dentro da transcrição fonológica, a identificação do segmento que ocorre logo antes e logo depois é previsível, bem como a estrutura da sílabas que contém a vogal.

A codificação automática da Vogal da Sílabas Seguinte foi possível devido à anotação sistemática, na trilha de vogais no Praat, de todas as vogais das sílabas seguintes aos segmentos-alvo; tanto a sua identificação (/a/, /i/ etc.) quanto as respectivas medidas de F1 e F2 se referem aos dados da linha seguinte da planilha. Por sua vez, a identificação da Vogal Tônica de cada item lexical é possível por

sua marcação na transcrição fonológica pelo símbolo (˘), realizada pelo *script* silac; uma vez identificada a sílaba tônica, dela se apagam as consoantes (p.ex., de /so-˘fr~e-do/, identifica-se a tônica /fr~e/ pela marca de tonicidade e a delimitação silábica pelos hífens, e dela se apagam as consoantes “f” e “r”). Por fim, a distância da vogal pretônica em relação à vogal tônica é possível por meio da contagem de hífens. Por exemplo, em “m<o>-˘raR”, há apenas um hífen entre a vogal pretônica < > e a marcação de tonicidade ˘, de modo que a distância da sílaba tônica é “1”.<sup>9</sup>

A codificação de variáveis sociais, por sua vez, também é previsível, já que as características de amostra, sexo, idade, idade de chegada etc. são constantes para cada falante. Tais informações para cada informante foram inseridas no *script* de codificação, que as repetiu em tantas linhas quanto fossem as ocorrências extraídas daquele arquivo. Além dessas seis variáveis sociais, incluíram-se a amostra (ALSP, PBSP e SP2010) e o informante para análises específicas.

Como variável dependente/resposta deste estudo, foram empregadas as medidas de F1 e F2 de cada vogal, tanto em seu formato bruto, extraído pelo *script* Vowel Analyzer, quanto em formato normalizado. Para normalização, empregou-se o método de Lobanov (LOBANOV, 1971), por meio da função *norm.lobanov* do pacote *vowels* (KENDALL; THOMAS, 2015) na plataforma R. A normalização de vogais permite maior comparabilidade dos espaços vocálicos entre diferentes indivíduos ao neutralizar diferenças de medidas de formantes decorrentes de características fisiológicas e de trato vocal (p.ex., mulheres tendem a apresentar medidas mais altas de formantes em comparação com as medidas dos homens para as mesmas vogais).

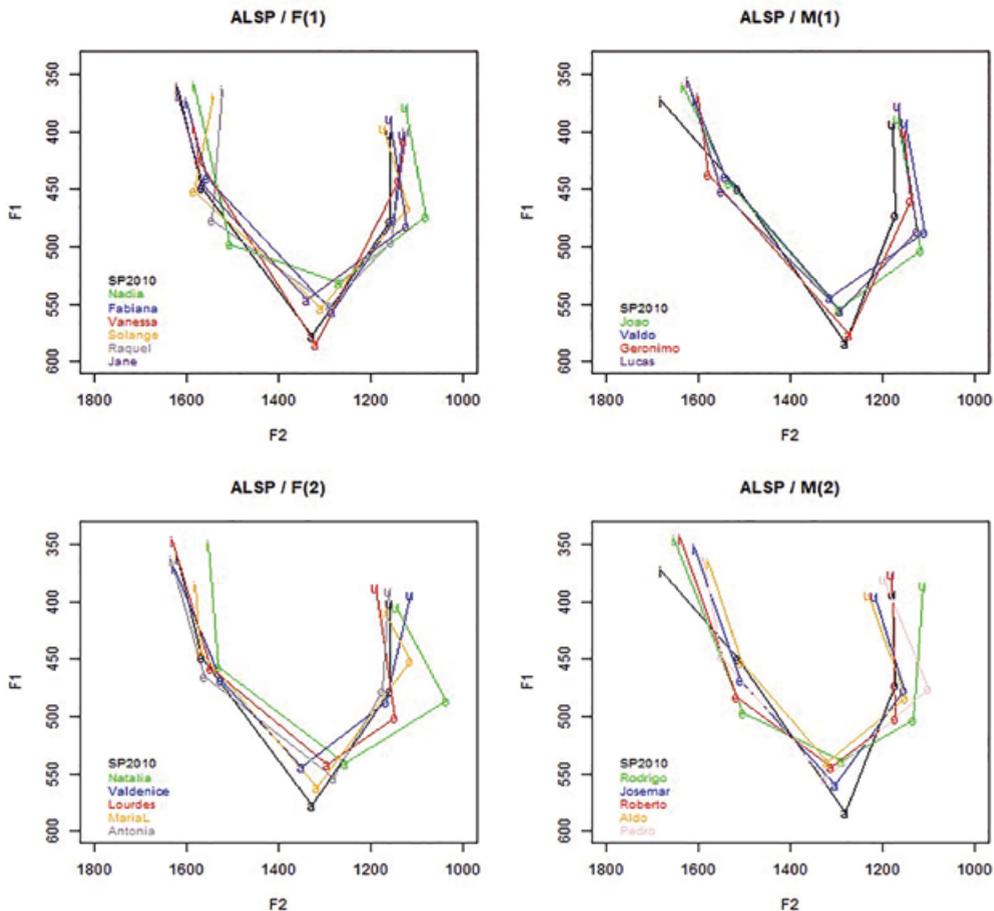
## RESULTADOS

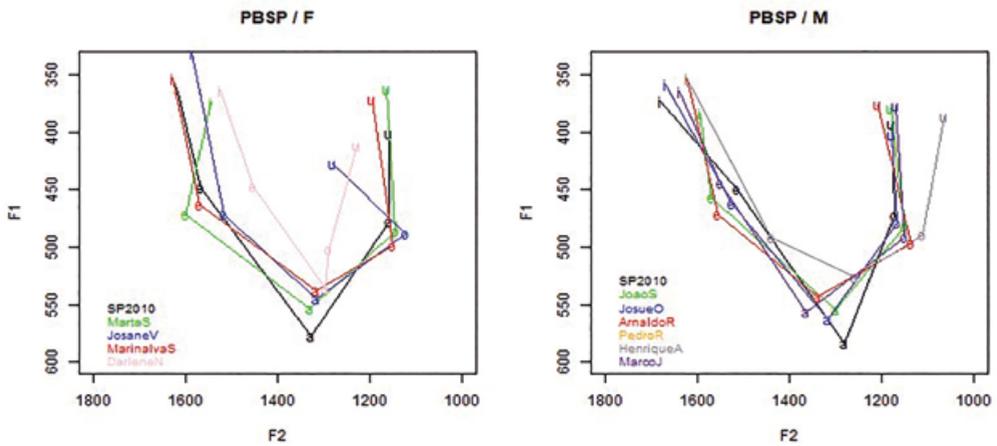
Na análise das vogais médias pretônicas, foram extraídas ocorrências apenas em contexto favorecedor de abaixamento vocálico, de acordo com a descrição de Pereira (1997) para a fala de João Pessoa-PB: vogais pretônicas em palavras cuja sílaba seguinte contém uma vogal [-alta] /ε, a, ɔ/ (como em *relógio*, *cortava*, mas não em *menino*) ou nasal /ĩ, ẽ, ã, õ, ù/ (como em *romã*, *diferente*). Tomou-se esta decisão pois é a realização relativamente mais baixa [ε]/[ɔ] da vogal pretônica, e não seu alicamento [i]/[u], que mais propriamente diferencia as variedades; inte-

<sup>9</sup> Uma vez extraídas e codificadas as informações referentes à vogal da sílaba seguinte e da sílaba tônica, as ocorrências correspondentes a vogais tônicas (identificadas pela ausência de parênteses na coluna “Vowel”) foram apagadas do arquivo de dados, já que não se referem a segmentos-alvo deste estudo.

ressa investigar quais fatores favorecem a realização relativamente menos baixa das vogais nos contextos em que tendem a ocorrer vogais baixas; tais usos sinalizam mais claramente a acomodação aos padrões da comunidade anfitriã do que, por exemplo, os dados de alçamento vocálico (viz. [mi.'ni.nu]), que ocorre tanto na comunidade de origem quanto na comunidade anfitriã. Para cada informante, foram codificados 50 dados de /e/ e 50 de /o/, acrescidos de 30 dados de cada uma das vogais /i/, /a/ e /u/ em posição pretônica, a fim de visualizar seus espaços vocálicos. A Figura 5 mostra tais espaços para cada falante, divididos por amostra (ALSP e PBSP), sexo (mulheres à esquerda e homens à direita) e faixa etária (para a amostra de alagoanos). Nesses gráficos, a altura média das vogais dos paulistanos (SP2010) são indicados pela linha preta, como uma referência da fala de nativos com a qual os migrantes estão em contato.

**Figura 5:** Espaços vocálicos pretônicos de migrantes alagoanos (ALSP) e paraibanos (PBSP) residentes em São Paulo.





Fonte: elaboração própria.

Na Figura 5, nota-se que a altura média das vogais /e/ e /o/ difere amplamente entre os falantes. Alguns deles apresentam vogais claramente mais baixas, com medidas maiores de F1 (i.e., realizações mais abertas) do que a dos paulistanos, como a vogal /e/ de Nádia e Raquel da amostra ALSP / F(1) (topo à esquerda) e a vogal /o/ de Roberto e Rodrigo (ALSP / M(2), centro à direita). Há, por outro lado, falantes cujas vogais médias são bastante semelhantes àquelas da comunidade anfitriã, como a vogal /e/ de Solange (ALSP, topo à esquerda) e a vogal /o/ de Josué (abaixo à direita). Além disso, percebe-se que a distância entre as vogais médias em relação às altas /i, u/ e a baixa /a/ para paulistanos tende a ser maior, ao passo que as vogais /e/ e /o/ dos migrantes se encontram relativamente mais próximas de suas vogais /a/. É possível que, além da altura das vogais médias pretônicas em Hz, a distância relativamente menor entre elas e a vogal baixa contribua para a percepção de que os migrantes têm vogais mais abertas.

Tendo em vista a ampla variabilidade entre indivíduos, cabe então avaliar se e quais variáveis sociais se correlacionam com a aproximação de padrões linguísticos da comunidade anfitriã, o que sinalizaria maior acomodação a essa variedade. As Tabelas 1 e 2 apresentam os resultados de análises de regressão linear em modelos de efeitos mistos (OUSHIRO, 2017), em que se compara a altura da vogal /e/ (F1) para diferentes fatores das variáveis previsoras sociais. Para cada uma das vogais, foram criados dois modelos lineares: o primeiro incluía as variáveis estratificadoras da amostra – Sexo/Gênero, Faixa Etária e Escolaridade –, e o segundo incluía outras variáveis de interesse a este estudo – Idade de Migração, Tempo de Residência em São Paulo e Motivo da Migração. Cada

conjunto de variáveis foi analisado em rodadas separadas pela falta de ortogonalidade entre elas, uma vez que as amostras haviam sido coletadas inicialmente para outros fins. Por exemplo, todos os falantes que vieram “para estudar” têm nível superior de escolaridade e são da primeira ou da segunda faixa etária, e todos os falantes que disseram ter vindo pela “família” são do sexo feminino, já que correspondem aos casos em que o companheiro havia migrado previamente e, depois de estabelecido, “mandou trazer a família”. Todos os modelos incluíram o Informante como uma variável aleatória, que permite avaliar se os padrões observados se devem de fato às variáveis sociais ou se podem ter sido enviesados pelo comportamento idiossincrático de um ou outro falante (OUSHIRO, 2017).<sup>10</sup>

**Tabela 1:** Estimativas para altura da vogal /e/ pretônica (medidas de F1 em Hz) em modelo de regressão linear de efeitos mistos para variáveis predictoras estratificadoras da amostra.

N = 1.916.

Previsoras	Estimativa	Erro padrão	Valor-t	p	
Intercepto	462,48	8,457	54,683	< 0,001	***
<b>Sexo</b>					
Feminino (valor de referência)					
Masculino	2,05	6,48	0,32	0,754	
<b>Escolaridade</b>					
Fundamental (valor de referência)					
Médio	-5,88	8,47	-0,69	0,494	
Superior	-9,40	9,47	-0,99	0,330	
<b>Faixa etária</b>					
20-34 (valor de referência)					
35-59	-3,52	7,37	-0,48	0,637	
60+	8,64	11,24	0,77	0,449	

Modelo: F1.NORM ~ SEXO + ESCOLARIDADE + FAIXA.ETARIA + (1|INFORMANTE), data = vogal.e. As estimativas dos valores de referência coincidem com a do Intercepto.

Fonte: Elaboração própria.

<sup>10</sup> Os modelos de regressão logística consideram apenas os dados dos migrantes, uma vez que as variáveis Idade de Migração, Tempo de Residência e Motivo da Migração não se aplicam aos paulistanos nativos.

**Tabela 2:** Estimativas para altura da vogal /e/ pretônica (medidas de F1 em Hz) em modelo de regressão linear de efeitos mistos para outras variáveis previsoras sociais. N = 1.916.

Previsoras	Estimativa	Erro padrão	Valor-t	p	
Intercepto	436,94	14,75	29,62	<0,001	***
<b>Idade de Migração</b>					
9 a 17 anos (valor de referência)					
18 a 24 anos	6,37	6,74	0,94	0,352	
25+ anos	17,44	8,22	2,12	0,044	*
<b>Tempo em SP</b>					
Menos de 10 anos (valor de referência)					
11-29 anos	-9,13	10,43	-0,88	0,392	
30+ anos	-9,06	11,31	-0,80	0,432	
<b>Motivo de migração</b>					
Estudo (valor de referência)					
Família	8,74	11,46	0,76	0,454	
Qualidade de vida	7,88	12,38	0,64	0,532	
Trabalho	13,18	10,15	1,30	0,209	

Modelo: IDADE.MIGRACAO + TEMPO.SP + MOTIVO.MIGRACAO + (1|INFORMANTE), data = vogal.e. As estimativas dos valores de referência coincidem com a do Intercepto.

Fonte: Elaboração própria.

Como se trata de modelo ainda pouco usual nos estudos sociolinguísticos brasileiros, cabem inicialmente algumas considerações sobre sua leitura e interpretação. As tabelas apresentam, em sua primeira coluna, os níveis das variáveis independentes fixas. O valor de Intercepto se refere à estimativa de todos os níveis de referência, que são estabelecidos pela plataforma R a partir de um critério alfabético; assim, para a variável Sexo/Gênero, que tem os níveis “feminino” e “masculino”, o programa toma como referência o primeiro, pois “f” precede “m”. Os demais níveis de referência são “20-34 anos” para Faixa Etária, “Fundamental” para Escolaridade etc. Desse modo, na Tabela 1, o Intercepto se refere à altura da vogal média pretônica /e/ na fala de mulheres entre 20 e 34 anos com nível fundamental de escolaridade, que foi de 462,48 Hz.

A segunda coluna apresenta as medidas estimadas para a variável resposta (valores de F1), em sua unidade de análise (Hz). Para cada um dos níveis além do nível de referência, a estimativa se refere à *diferença* entre o Intercepto e o

respectivo fator. Por exemplo, a estimativa da altura da vogal /e/ para homens é 2,05 Hz acima (pelo valor positivo) da estimativa para as mulheres; a estimativa para falantes entre 35 e 59 anos é 3,51 Hz abaixo (pelo valor negativo) da estimativa para falantes entre 20 e 34 anos; a estimativa para falantes com 60 anos ou mais é 8,64 Hz acima da estimativa para aqueles entre 20 e 34 anos. Para saber a verdadeira estimativa para determinado fator que não se encontra no Intercepto, é necessário somar as estimativas do Intercepto e do fator. Por exemplo, para os falantes do sexo masculino, a estimativa de altura da vogal /e/ pretônica é de  $462,48 + 2,05 = 464,53$  Hz. Apesar de parecer pouco intuitiva a apresentação de resultados na forma de *diferenças* em relação ao Intercepto (em vez dos valores estimados em si), ela permite verificar mais rapidamente o quanto tal valor difere de zero: se a estimativa for nula (ou próxima dela), isso significa que é praticamente idêntica à do Intercepto; ou, visto de outro modo, que provavelmente não há diferença significativa entre os níveis da variável.

A terceira coluna apresenta o erro padrão, uma medida de dispersão que dá indícios de quão precisa é a estimativa. Quanto maior o valor de erro padrão, maior é a variabilidade nas medições. O valor-*t*, na quarta coluna, é calculado pela razão Estimativa / Erro Padrão (p.ex., para masculino:  $2,05 / 6,8 = 0,32$ ). Tal valor é usado para calcular um valor de significância, que mede a probabilidade de se observar tal distribuição em caso de a hipótese nula ser verdadeira; por convenção, estabelece-se um limite de até 5% ( $p < 0,05$ ) para que uma diferença seja considerada significativa, e representam-se os níveis de significância por meio de \* $p < 0,05$ , \*\* $p < 0,01$ , e \*\*\* $p < 0,001$ . Portanto, as linhas que apresentam asteriscos representam níveis em que se verificam estimativas significativamente diferentes em relação ao Intercepto e, portanto, correlação significativa entre a variável resposta (nesta análise, a altura da vogal pretônica /e/) e a respectiva variável previsora.

Nas Tabelas 1 e 2, verifica-se que, das seis variáveis sociais, apenas Idade de Migração apresenta correlação com a altura da vogal pretônica /e/: embora não haja diferença significativa entre os falantes que migraram até os 17 anos de idade e aqueles que migraram entre 18 e 24 anos ( $p = 0,35$ ), há diferença significativa entre os que se deslocaram mais cedo e aqueles que migraram com mais de 30 anos ( $p = 0,04$ ). A estimativa para estes indivíduos, 17,44 Hz, indica que a altura da vogal /e/ para eles é mais baixa, uma vez que o valor de F1 é mais alto (Intercepto + 17,44 Hz). Assim, quanto mais tarde migrou o falante, mais baixas tendem a ser suas vogais /e/ - ou, visto de outra perspectiva, quanto mais cedo migrou o falante, mais altas tendem a ser suas vogais /e/, ou seja, mais próximas

do padrão paulistano. A falta de correlação significativa com as demais variáveis sociais é indício de que a variação entre os indivíduos é mais importante para entender a variação na altura da vogal /e/ dos migrantes do que fatores como Sexo/ Gênero, Escolaridade e Tempo em São Paulo, e que tais variáveis previsoras não promovem a aproximação da fala dos migrantes aos padrões paulistanos para esta variável sociolinguística.

As Tabelas 3 e 4 abaixo contêm os resultados para modelos de efeitos mistos em que se testaram correlações entre as variáveis previsoras sociais e a altura da vogal pretônica /o/.

**Tabela 3:** Estimativas para altura da vogal /o/ pretônica (medidas F1 em Hz) em modelo de regressão linear de efeitos mistos para variáveis previsoras estratificadoras da amostra.

N = 1.645.

Previsoras	Estimativa	Erro padrão	Valor-t	p	
Intercepto	481,70	7,92	60,82	< 0,001	***
<b>Sexo</b>					
Feminino (valor de referência)					
Masculino	10,69	5,94	1,80	0,083	
<b>Escolaridade</b>					
Fundamental (valor de referência)					
Médio	-8,60	7,84	-1,10	0,282	
Superior	-11,62	8,65	-1,34	0,191	
<b>Faixa etária</b>					
20-34 (valor de referência)					
35-59	3,63	6,79	0,53	0,598	
60+	5,69	10,48	0,54	0,591	

Modelo: F1.NORM ~ SEXO + ESCOLARIDADE + FAIXA.ETARIA + (1|INFORMANTE), data = vogal.o. As estimativas dos valores de referência coincidem com a do Intercepto.

Fonte: Elaboração própria.

**Tabela 4:** Estimativas para altura da vogal /o/ pretônica (medidas F1 em Hz) em modelo de regressão linear de efeitos mistos para outras variáveis predictoras sociais. N = 1.645.

Previsoras	Estimativa	Erro padrão	Valor-t	p	
Intercepto	479,88	12,67	37,86	< 0,001	***
<b>Idade de Migração</b>					
9 a 17 anos (valor de referência)					
18 a 24 anos	12,94	6,01	2,15	0,04	*
25+ anos	9,49	7,51	1,26	0,22	
<b>Tempo em SP</b>					
Menos de 10 anos (valor de referência)					
11-29 anos	3,42	8,40	0,41	0,688	
30+ anos	-5,71	9,05	-0,63	0,535	
<b>Motivo de migração</b>					
Estudo (valor de referência)					
Família	-4,55	9,66	-0,47	0,642	
Qualidade de vida	0,75	10,33	0,07	0,943	
Trabalho	0,98	8,35	0,12	0,908	

Modelo: IDADE.MIGRACAO + TEMPO.SP + MOTIVO.MIGRACAO + (1|INFORMANTE), data = vogal.o. As estimativas dos valores de referência coincidem com a do Intercepto.

Fonte: Elaboração própria.

De modo semelhante aos resultados para a vogal /e/, verifica-se correlação apenas com a variável Idade de Migração; neste caso, há diferença significativa entre os falantes que migraram entre 18 e 24 anos (+12,94 Hz,  $p = 0,04$ ) em relação aos que chegaram a São Paulo mais cedo (até os 17 anos). Os falantes que chegaram após os 30 anos não apresentam diferença significativa em relação aos mais novos, embora a estimativa também indique vogais relativamente mais baixas (i.e., medidas de F1 mais altas: +9,49 Hz) para eles. Nenhuma das demais variáveis sociais apresenta correlação significativa com a altura da vogal /o/, o que indicia novamente a preponderância da variável Idade de Migração para a aquisição de padrões paulistanos -- quanto mais cedo chegou, maior tendência à aquisição de vogais relativamente menos baixas.<sup>11</sup>

<sup>11</sup> Para um modelo e uma análise pormenorizada do padrão dos indivíduos, ver Oushiro (2019).

## DISCUSSÃO

A Figura 5 mais acima mostra a ampla variação no espaço vocálico entre os indivíduos migrantes, que se diferenciam ou se aproximam do padrão vocálico pretônico da comunidade anfitriã de modo gradual. Não é o caso de haver apenas dois grupos de falantes, os que adquiriram os padrões paulistanos e os que não adquiriram; a altura das vogais /e/ e /o/ se organiza em um contínuo, desde vogais mais abertas (i.e., mais próximas do padrão da comunidade de origem) até aquelas que se aproximam e não se diferenciam estatisticamente do padrão paulitano. Ora, é certo que a variação linguística já é prevista por qualquer estudo sociolinguístico; no entanto, o tratamento dos dados das vogais médias pretônicas como uma variável contínua permite visualizar e captar não apenas a variação, mas também a gradação.

Ao mesmo tempo, a ampla variação entre os indivíduos faz com que seja imprescindível analisar os dados em um modelo que permita incluir os falantes como uma variável, simultaneamente às demais variáveis previsoras.<sup>12</sup> A análise de efeitos mistos se caracteriza pela inclusão de variáveis fixas (que podem ser replicadas em diferentes estudos) e variáveis aleatórias (próprias da amostra sob análise), de modo a obter estimativas mais confiáveis dos padrões generalizáveis a toda a comunidade, uma vez que o efeito de fatores idiossincráticos (como o comportamento de diferentes falantes) pode ser controlado.

Ao controlar efeitos idiossincráticos, nota-se que, das seis variáveis sociais analisadas, apenas a Idade de Migração do falante se correlaciona significativamente com a altura das vogais médias pretônicas. Tanto para a vogal /e/ quanto para a vogal /o/, observou-se que os alagoanos e paraibanos que chegaram mais cedo a São Paulo (até os 17 anos) realizam suas vogais de modo significativamente menos baixo comparativamente àqueles que migraram mais tarde. Tal resultado está de acordo com a expectativa de que os mais jovens tendem a adquirir uma nova língua e um novo dialeto mais facilmente.

Por outro lado, outras variáveis que plausivelmente teriam efeito sobre a altura das vogais não se mostraram correlacionadas (ver CHAMBERS, 1992; TRUDGILL, 1986). No senso comum (inclusive de participantes desta pesquisa), esperar-se-ia que a permanência na nova comunidade por longo período de

---

<sup>12</sup> A análise dos padrões de indivíduos também é possível no programa GoldVarb X, codificando-se uma variável que identifique cada falante. Contudo, nesse caso, é necessário excluir todas as demais variáveis sociais, uma vez que não serão ortogonais ao falante, pois um mesmo indivíduo não pode ser do sexo masculino e feminino, de diferentes faixas etárias etc.

tempo também conduziria à assimilação de seus traços. A falta de correlação com a variável Tempo de Residência em São Paulo contradiz essa hipótese: pelo menos para as vogais médias pretônicas, que não é caso. Tampouco há diferenças quanto aos perfis de migrantes, que se deslocaram por motivos diversos: para estudar, para trabalhar, por conta da família ou para uma melhor qualidade de vida. Em estudo pioneiro sobre as atitudes de migrantes baianos e pernambucanos em São Paulo, Alves (1979) havia levantado a hipótese de que os falantes de “nível sociocultural baixo” tenderiam a assimilar mais as formas linguísticas da comunidade anfitriã, em comparação com os de “nível sociocultural alto”, pelo fato de os primeiros sofrerem mais preconceito devido à sua condição socioeconômica, o que os motivaria a “se camuflar linguisticamente”. A presente análise não invalida essa hipótese, mas aponta para a necessidade de uma tipologia mais ampla do que ocorre a diferentes variáveis sociolinguísticas em situação de contato dialetal: seria o caso de que algumas variáveis são mais suscetíveis à assimilação do que outras?

Além disso, não se verificou correlação significativa entre a altura das vogais pretônicas e variáveis sociais tradicionalmente analisadas nos estudos sociolinguísticos: o Sexo/Gênero, a Faixa Etária e o Nível de Escolaridade dos falantes. Nesse caso, cabe reavaliar as expectativas e as hipóteses de acordo com a comunidade específica sob análise – constatação que se aplica, evidentemente, a qualquer estudo. Como visto, as amostras desta pesquisa haviam sido coletadas previamente para outros fins, mas, face aos resultados aqui apresentados, estudos futuros sobre a fala de migrantes devem buscar analisar mais sistematicamente o papel de outras variáveis mais relevantes para explicar seus padrões de fala – como Idade e Motivo de Migração, Tempo de Residência na nova comunidade, Redes Sociais, Identidades, entre outras.

## CONSIDERAÇÕES FINAIS

A análise sobre a fala de 32 migrantes nordestinos que residem em São Paulo mostrou que os padrões desses falantes se distribuem em um contínuo de aproximação–distanciamento daqueles da comunidade anfitriã e que, dentre seis variáveis sociais analisadas, apenas a Idade de Migração se correlaciona significativamente com a variação na altura das vogais médias pretônicas.

Paralelamente ao objetivo de descrever os padrões de acomodação na fala de migrantes alagoanos e paraibanos em São Paulo, propôs-se neste trabalho descrever e avaliar os métodos empregados na identificação, extração e preparo dos dados, tarefas que antecedem a análise estatística. Ressalta-se o uso de

ferramentas como o R, o ELAN e o Praat para a automação e agilização de tais tarefas, no tratamento de milhares de dados de uma variável fonética.

Partindo-se da hipótese de que a acomodação dialetal por parte de migrantes alagoanos e paraibanos numa capital do sudeste quanto à altura de vogais pretônicas constitui um processo gradual, interessou a esta pesquisa analisar a variável de modo contínuo – as medidas de F1 (e, auxiliarmente, de F2) e os respectivos espaços vocálicos dos migrantes em contraste com os de falantes nativos da comunidade anfitriã, em vez de categorizações discretas em “vogais altas”, “vogais médias” e “vogais média-baixas”. Trata-se de um percurso de pesquisa que muitas vezes não é seguido em estudos variacionistas; não raro se subordinam as questões de pesquisa à aplicação automática de um conjunto restrito de métodos, em vez de se buscarem novas ferramentas que permitam o pesquisador analisar fenômenos linguísticos sob nova perspectiva.

O desenvolvimento dos *scripts* silac, silacpret e o de codificação automática de variáveis independentes, assim como a testagem da aplicação de ferramentas já existentes como o EasyAlign e o Vowel Analyzer, não constituem tarefas triviais, mas o esforço sistemático de buscar métodos mais eficientes e adequados para a análise de dados. A dedicação desse tempo à elaboração de procedimentos metodológicos certamente é recompensada pela perspectiva de sua aplicação em estudos futuros. O transcritor fonológico *silac* (OUSHIRO, 2018) pode ser empregado em diversos estudos sobre variáveis fonéticas, não apenas sobre vogais pretônicas. O emprego de *scripts* no Praat abre o caminho para análises acústicas mais refinadas, que não dependem apenas da percepção auditiva do pesquisador, assim como para o processamento de uma grande quantidade de dados. Os procedimentos e *scripts* para o tratamento de vogais podem ser aplicados tanto na expansão das amostras a serem analisadas na presente pesquisa, quanto em outros estudos.

Espera-se, por fim, que este trabalho contribua para a reflexão continuada no campo de estudos sobre a necessidade de questões e métodos caminharem lado a lado: novas questões exigem novos métodos, e novos métodos abrem a possibilidade de novos questionamentos.

## REFERÊNCIAS

ALVES, Maria Isolete Pacheco Menezes. *Atitudes linguísticas de nordestinos em São Paulo: abordagem prévia*. Dissertação de Mestrado. Campinas: IEL/UNICAMP, 1979.

BATTISTI, Elisa; DORNELLES FILHO, Adalberto Ayjara. Palatalização das plosivas alveolares em Flores da Cunha (RS): variação linguística e práticas sociais. *Alfa* vol. 56, n.3, 1117-1149, 2012.

BOERSMA, Paul; WEENINK, David. Praat: doing phonetics by computer, 2018. Disponível em: <<http://www.fon.hum.uva.nl/praat/>>. Acesso em: 26 jan. 2019.

BORTONI, Stela M.; GOMES, Christina A; MALVAR, Elisabete. A variação das vogais médias pretônicas no português de Brasília: um fenômeno neogramático ou de difusão lexical? *Revista Estudos Linguísticos*, vol. 1, 9-29, 1992.

BRESCANCINI, Claudia; MONARETTO, Valéria Neto de Oliveira. Os róticos no sul do Brasil: panorama e generalizações. *Signum: Estudos Linguísticos*, vol. 11, n. 2, p. 51–66, 2008.

BRANDÃO, Silvia Figueiredo. Variação em coda silábica na fala popular fluminense. *Revista da ABRALIN*, vol. 7, n. 1, 177-189, jan./jun. 2008.

CALLOU, Dinah; MORAES, João; LEITE, Yonne. Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In: *Gramática do português falado*, vol. VI. Campinas: Editora da Unicamp, 1996.

CHAMBERS, Jack K. Dialect acquisition. *Language*, vol. 68, n. 4, p. 673-705, 1992.

HELLWIG, B.; GEERTS, J. ELAN – Linguistic Annotator. Versão 5.4. Disponível em: <<http://www.mpi.nl/corpus/manuals/manual-elan.pdf>>, 2018. Acesso em: 26 jan. 2019.

GOLDMAN, Jean-Philippe. *EasyAlign*: an automatic phonetic alignment tool under Praat. In: Proceedings of InterSpeech, 2011. Disponível em: <<https://archive-ouverte.unige.ch/unige:18188>>. Acesso em: 30 set. 2016.

GOMES DA SILVA, Fernando. *Alagoanos em São Paulo e a concordância nominal de número*. Dissertação de Mestrado. São Paulo: FFLCH/USP, 2014. 103f.

HORA, Dermeval da; NEGRÃO, Esmeralda V. (eds.). *Estudos da Linguagem: casamento entre temas e perspectivas*. João Pessoa: Ideia, 2011.

KENDALL, Tyler; THOMAS, Erik R. Package ‘vowels’, 2015. Pacote para a plataforma R. Disponível em: <<http://blogs.uoregon.edu/vowels/2012/11/08/vowels-r-1-2/>>. Acesso em: 30 set. 2016.

LABOV, William. *Padrões sociolinguísticos*. São Paulo: Editora Parábola, 2008 [1972].

LOBANOV, Boris M. Classification of Russian vowels spoken by different speakers. *Journal of Acoustic Society of America*, vol. 49, n. 2, 606-608, 1971.

MACEDO, Sandra Siqueira de. A palatalização de /s/ em coda silábica no falar culto recifense. Dissertação de Mestrado. Recife: UFPE, 2004. 100f.

MARTINS, Mariana de Souza. *A palatalização de oclusivas dentais em contato dialetal*. Dissertação de Mestrado. Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2008. 146f.

MENDES, Ronald Beline; OUSHIRO, Livia. O paulistano no mapa sociolinguístico brasileiro. *Alfa*, vol. 56, n. 3, 973–1001, 2012.

OLIVEIRA, Josane M. *O apagamento do /R/ implosivo na norma culta de Salvador*. Dissertação de Mestrado. Salvador: UFBA, 1999. 80 f.

OUSHIRO, Livia. *Uma análise variacionista para as Interrogativas-Q*. Dissertação de Mestrado. São Paulo: FFLCH/USP, 2011. 174f.

OUSHIRO, Livia. Transcrição de entrevistas sociolinguísticas com o ELAN. In: FREITAG, Raquel M. Ko. *Metodologia de Coleta e Manipulação de Dados em Sociolinguística*. São Paulo: Blucher, 2014, p. 117-132. Disponível em: <<http://openaccess.blucher.com.br/article-details/transcrio-de-entrevistas-sociolinguisticas-com-o-elan-18959>>. Acesso em: 30 set. 2016.

OUSHIRO, Livia. Processos de acomodação dialetal na fala de nordestinos residentes em São Paulo. Projeto Regular FAPESP (2016/04960-7). 2016. Ms.

OUSHIRO, Livia. *Introdução à estatística para linguistas*, v. 1.0.1 (dez/2017). DOI 10.5281/zenodo.822069. Disponível em: <<https://zenodo.org/record/1202201>>. Acesso em: 29 jan. 2019.

OUSHIRO, Livia. silac: *Transcritor fonológico do português*. 2018. Versão online (v0.5.1). Disponível em: <[oushiro.shinyapps.io/silac](http://oushiro.shinyapps.io/silac)>. Acesso em: 18 jan. 2019.

OUSHIRO, Livia. Linguistic uniformity in the speech of Brazilian internal migrants in a dialect contact situation. In: CALHOUN, Sasha; ESCUDERO, Paola; TABAIN, Marija; WARREN, Paul (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*, Melbourne, Australia 2019. Disponível em: <<https://icphs2019.org/icphs2019-fullpapers/>>. Acesso em: 10 set. 2019.

PEREIRA, Regina C. M. *As vogais médias pretônicas na fala pessoense urbana*. Dissertação de Mestrado. João Pessoa: UFPB, 1997.

R CORE TEAM. R: *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2018. Disponível em: <<https://www.r-project.org/>>. Acesso em: 26 jan. 2019.

RIEBOLD, John. *Vowel analyzer*. Ms., 2013. (script do Praat). Disponível em: <<https://raw.githubusercontent.com/jmriebold/Praat-Tools/master/Vowel-Analyzer.praat>>. Acesso em: 30 set. 2016.

SANKOFF, David; TAGLIAMONTE, Sali; SMITH, Eric. Goldvarb X: A variable rule application for Macintosh and Windows. Disponível em: <<http://individual.utoronto.ca/tagliamonte/goldvarb.html>>. Acesso em: 19 jan. 2019.

TAGLIAMONTE, Sali. *Variationist sociolinguistics: change, observation, interpretation*. Malden, MA: Wiley-Blackwell, 2012.

TRUDGILL, Peter. *Dialects in contact*. New York: Basil-Blackwell, 1986.

YACOVENCO, Lilian. *As vogais médias pretônicas no falar culto carioca*. Tese de Doutorado, Rio de Janeiro: UFRJ, 1993. 193f.



## QUESTIONS AND METHODS PRETONIC MIDVOWELS IN THE SPEECH OF NORTHEASTERN MIGRANTS IN DIALECTAL CONTACT

*Livia Oushiro (Universidade Estadual de Campinas)*

This article has two objectives: (i) to present the methods employed in an analysis of the variable pretonic midvowel height in the speech of 32 migrants from the Northeastern states of Alagoas and Paraíba living in the Southeastern state of São Paulo/Brazil, in a dialect contact situation; and (ii) to call attention to two limitations of variationist studies, namely, the great amount of time spent in identifying, extracting, and coding data, and the subordination of research questions to a restricted set of methods. Sociolinguistic studies of the variable pretonic midvowel height have generally treated it categorically, considering the factors “high” (vowels [i, u]), “mid-high” (vowels [e, o]), and “mid-low” (vowels [ɛ, ɔ]), analyzed in different binary models in Varbrul. From an initial expectation that migrants’ speech is better characterized as a continuum, exhibiting patterns which are possibly intermediary between that of the native and the host community, and that their speech could present more individual variation than in the speech of native/prototypical speakers, we analyzed their vowels in mixed effects linear regression models with vowel height (measured in F1) as response variable, speaker as a random predictor, and other linguistic and social predictors as fixed effects – namely, preceding and following phonological context, height of following syllable vowel, stressed vowel, distance from stressed vowel,

structure of pretonic syllable, sex/gender, age group, level of education, age of arrival, length of residence, and motivation for migration). An additional sample of 7 native São Paulo speakers was also analyzed as a reference group.

The sociolinguistic interviews were first transcribed and time-aligned in ELAN (HELLWIG; GEERTS, 2018). The orthographic transcript was transformed into a phonological transcription through the use of *silac* (OUSHIRO, 2018), an R code available as an online app which outputs a transcript containing word syllabification and stress marking. The transcripts were then imported into Praat (BOERSMA; WEENINK, 2018), in which the Easy-Align plug-in (GOLDMAN, 2011) automatically segmented the wave, facilitating the identification of mid-vowels starting and ending points. After that, the Praat script Vowel Analyzer (RIEBOLD, 2013) extracted, into a semi-coded spreadsheet, F1 and F2 measurements at three points of around 14,000 vowels, which included all five Brazilian Portuguese pretonic vowels and their respective following syllable vowels. Finally, linguistic and social predictors were automatically coded and pretonic midvowels were Lobanov-normalized (LOBANOV, 1971) through an R script using the package vowels (KENDALL; THOMAS, 2015).

Analyses of six social predictors show that only migrants' age of arrival in the new community correlate significantly with pretonic midvowel height: the earlier the arrival, the greater the approximation to the host community's pattern, both for vowel /e/ and /o/. Differently from expected, predictors such as length of residence, motivation for migration, and sex/gender didn't show significant correlations, which suggests the need for more systematic analyses of a number of sociolinguistic variables in order to propose a typology of variables in dialect contact situation.

The methods herein developed and applied allowed for a more refined acoustic analysis of a great amount of migrants' speech data and for the observation of the effect of individual and social constraints in dialect contact and acquisition. They are all applicable to the study of other variables. These methods are not yet regularly employed in Brazilian sociolinguistic studies and have not had much publishing space, which may be symptomatic of a dependence on a single statistical method which, in turn, restricts the scope of research questions in the field. We argue for the need of developing and publishing methods (as well as results) and for the continued reflection upon research questions and methods: new questions require new methods, and new methods may lead to new insights.

## REFERENCES

- BOERSMA, Paul; WEENINK, David. Praat: doing phonetics by computer, 2018. Disponível em: <<http://www.fon.hum.uva.nl/praat/>>. Acesso em: 26 jan. 2019.
- HELLWIG, B.; GEERTS, J. ELAN – Linguistic Annotator. Versão 5.4. Disponível em: <<http://www.mpi.nl/corpus/manuals/manual-elan.pdf>>, 2018. Acesso em: 26 jan. 2019.
- GOLDMAN, Jean-Philippe. *EasyAlign*: an automatic phonetic alignment tool under Praat. In: Proceedings of InterSpeech, 2011. Disponível em: <<https://archive-ouverte.unige.ch/unige:18188>>. Acesso em: 30 set. 2016.
- KENDALL, Tyler; THOMAS, Erik R. Package ‘vowels’, 2015. Pacote para a plataforma R. Disponível em: <<http://blogs.uoregon.edu/vowels/2012/11/08/vowels-r-1-2/>>. Acesso em: 30 set. 2016.
- LOBANOV, Boris M. Classification of Russian vowels spoken by different speakers. *Journal of Acoustic Society of America*, vol. 49, n. 2, 606-608, 1971.
- OUSHIRO, Livia. silac: *Transcritor fonológico do português*. 2018. Versão online (v0.5.1). Disponível em: <[oushiro.shinyapps.io/silac](http://oushiro.shinyapps.io/silac)>. Acesso em: 18 jan. 2019.
- RIEBOLD, John. *Vowel analyzer*. Ms., 2013. (script do Praat). Disponível em: <<https://raw.githubusercontent.com/jmriebold/Praat-Tools/master/Vowel-Analyzer.praat>>. Acesso em: 30 set. 2016.



# O APAGAMENTO DA VOGAL POSTÔNICA MEDIAL EM DUAS VARIEDADES AFRICANAS DO PORTUGUÊS UMA DESCRIÇÃO PRELIMINAR

*Danielle Kely Gomes (UFRJ)*

### RESUMO

O apagamento da vogal postônica medial, processo que culmina a regularização de proparoxítonos ao padrão paroxítono, é um fenômeno antigo em Português, com raízes no latim. A redução fonética de vocábulos proparoxítonos é observada em diversas normas do Português Brasileiro (PB). Contudo, em outras realidades de uso do Português, sejam contextos em que este tem status de língua materna da maior parte dos indivíduos, sejam contextos em que assume feições de uma segunda língua (ou até mesmo de uma língua estrangeira), as descrições são escassas, ou mesmo inexistentes. Neste trabalho, propõe-se uma comparação entre dados do Português de São Tomé (PST) e do Português de Moçambique (PM), produzidos por doze informantes (seis em cada comunidade), pertencentes à faixa etária mais jovem (indivíduos com idades entre 18 e 35 anos). A hipótese que norteia este trabalho é a de que as proparoxítonas, padrão acentual atípico até para falantes de Português como L1, seriam consistentemente regularizadas a paroxítonas nas duas variedades, como efeito do contato do Português com as outras línguas que com ele coexistem em ambas as comunidades. Os resultados revelaram um comportamento similar entre as normas de uso analisadas no que

tange aos índices percentuais para a ocorrência do fenômeno: índices expressivos de apagamento da vogal átona medial (38% na variedade são tomense e 36%, nos dados da variedade moçambicana). Observaram-se similaridades na atuação das variáveis linguísticas que controlam o efeito dos segmentos adjacentes à postônica medial e divergências nos condicionamentos sociais que medem o efeito dos anos de escolarização e a influência do contato linguístico no apagamento da vogal postônica medial.

## INTRODUÇÃO

Em Português, verifica-se, no contexto postônico medial, a aplicação de duas regras fonológicas. Ao lado da regra de alteamento das vogais médias, que promove a realização fonética como vogal alta para os fonemas /e/ e /o/ (hipót[e] se~hipót[i]se; mét[o]du ~mét[u]du), a regra de apagamento também é produtiva (árv[o]re ~arv[u]re~arvri/arvi; cócegas ~ cóc[i]gas~coska), e regulariza itens proparoxítonos ao padrão acentual paroxítono. Processos que afetam o vocalismo postônico medial se fazem presentes desde o latim, e se conservam em Português, em sua diversidade.

Nesta investigação, apresentam-se os resultados de uma análise preliminar que compara duas variedades africanas do Português – a são-tomense (PST) e a moçambicana (PM) – no que se refere ao comportamento do vocalismo átono postônico medial, sobretudo ao processo de apagamento dessa vogal.

Para tanto, o capítulo contém as seguintes seções, na seguinte sequência: apresentam-se sínteses de trabalhos que focalizam o comportamento variável da postônica medial em Português; justifica-se a pertinência da pesquisa em andamento; discutem-se questões vinculadas a aspectos históricos e sociolinguísticos das comunidades investigadas; arrolam-se as hipóteses e a metodologia adotada para o tratamento dos dados; discutem-se os resultados e, por fim, tecem-se as considerações finais sobre o comportamento variável das proparoxítonas nas variedades são tomense e moçambicana.

## AS POSTÔNICAS MEDIAIS EM PORTUGUÊS: O QUE OS ESTUDOS JÁ MOSTRARAM

No âmbito do Português do Brasil (PB), um volume considerável de pesquisas registra a vitalidade e a persistência do fenômeno de apagamento da vogal postônica medial (cf., por exemplo, CAIXETA, 1989; AMARAL, 2000; SILVA, 2006, 2010; FONSECA, 2007; LIMA, 2008, 2017; RAMOS, 2009; CHAVES, 2011; GOMES, 2012). Os estudos, ainda que apresentem índices percentuais

diferenciados de aplicação da regra para cada variedade analisada, revelam consistência na atuação dos condicionamentos linguísticos e sociais, sobretudo a influência de restrições como a natureza dos contextos antecedente e subsequente à vogal, o traço de articulação da vogal apagada, o número de sílabas do vocábulo proparoxítono e a escolaridade.

No âmbito do Português Europeu (PE), os estudos são mais escassos (cf. os trabalhos de FERNANDES, 2007; GOMES, 2012, 2015), e revelam que a regularização de proparoxítonos a paroxítonos está condicionada a aspectos da configuração do sistema vocálico da variedade europeia, que propicia – de forma generalizada por todo o vocalismo átono – a atuação mais frequente de regras de apagamento.

No que concerne às variedades africanas do Português, o quadro ainda está por se revelar. Gomes (2017), em um estudo contrastivo entre a variedade brasileira, a europeia (com dados do *corpus* Concordância) e a são tomense (com dados da amostra *Variedades do Português<sup>1</sup>* - VAPOR), demonstra que há diferenças quantitativas consideráveis entre as variedades: no Português de São Tomé (PST), os índices de apagamento são elevados (34,7%, contra 10,8% para o PE e 2,6% para o PB), ainda que as variáveis estruturais para a implementação do apagamento atuem de forma semelhante nos três subconjuntos de dados analisados.

Do ponto de vista das restrições linguísticas, as três variedades se mostram sensíveis quanto à atuação dos contextos precedente e subsequente à vogal postônica medial: a possibilidade de ressilabificação das consoantes adjacentes à vogal – seja em direção à coda da sílaba tônica (cócegas > cosca), seja em direção ao ataque da sílaba átona final (fósforo > fosfro) – é o condicionamento decisivo para a implementação da regra no PB, no PE e no PST, embora as diferenças quantitativas sejam marcantes.

## A PROPOSTA DESTE TRABALHO: A COMPARAÇÃO PRELIMINAR ENTRE DUAS VARIEDADES AFRICANAS DO PORTUGUÊS

Gomes (2018) demonstra que o apagamento das vogais mediais no Português de São Tomé não pode estar desassociado de questões relativas ao contato linguístico. A relação que se estabelece entre o Português e o Forro, crioulo de base lexical portuguesa que coexiste com o Português na cidade de São Tomé,

---

<sup>1</sup> Corpus organizado por Tjerk Hagemeijer e recolhido no ano 2009, na cidade de São Tomé. O projeto está sediado no Centro de Linguística da Universidade de Lisboa.

local de recolha dos inquéritos que compõem o corpus *Variedades do Português*, se revela como uma restrição de relevância para a ocorrência do apagamento das vogais átonas mediais.

No trabalho citado, a autora demonstra uma correlação direta entre a queda da postônica medial e a frequência de uso do Forro: os indivíduos que afirmam usar eventualmente o crioulo tendem a apresentar índices mais elevados de apagamento de vogais, em conformidade a uma tendência do Forro em eliminar segmentos átonos, de forma a regularizar as palavras de origem portuguesa à estrutura silábica CVCV (FERRAZ, 1979, p. 47). Sendo assim, o comportamento da variedade são tomense é regulado por questões derivadas do multilinguismo generalizado que marca o contexto insular.

Para verificar se esse mesmo comportamento se mantém em outras variedades africanas do Português igualmente marcadas por contatos linguísticos massivos e consistentes, propõe-se, neste trabalho, uma comparação entre os resultados verificados para os seis informantes da primeira faixa etária de São Tomé e os seis informantes de Moçambique com o mesmo perfil. O *corpus* relativo ao Português de Moçambique (PM) pertence ao projeto *Estudo comparado dos padrões de concordância em variedades Africanas, Brasileiras e Europeias do Português*.<sup>2</sup> Os motivos que levam à comparação entre São Tomé e Moçambique e à caracterização dos *corpora* serão apresentados a seguir.

## PERFIL E HISTÓRICO SOCIOLINGÜÍSTICOS DAS COMUNIDADES ANALISADAS

A história de colonização de São Tomé pode ser dividida em dois ciclos. O primeiro, que vai do início da ocupação efetiva das ilhas (1493) até os fins do século XVI, é marcado pelo plantio da cana-de-açúcar. O segundo, a partir da segunda metade do século XIX, é caracterizado pelas culturas de café e cacau.

Do ponto de vista linguístico, no primeiro ciclo surge um *pidgin* que assegura a comunicação mínima entre os portugueses e os africanos que habitavam a ilha. Foi a partir da nativização desse *pidgin* que se originou um crioulo de base lexical portuguesa, a língua da comunidade de escravizados. De acordo com Gonçalves e Hagemeijer (2015, p. 88), o Forro – crioulo majoritário de São Tomé e Príncipe – é a continuação no tempo dessa protolíngua. A cultura açucareira de São Tomé e Príncipe entra em declínio nos fins do século XVI, com a inserção no cenário internacional do açúcar produzido no nordeste do Brasil. São Tomé

---

<sup>2</sup> Os inquéritos foram recolhidos por Silvia Rodrigues Vieira e Karen Cristina da Silvia Pisurno em setembro de 2016 na cidade de Maputo.

deixa de ser uma colônia de produção açucareira e se torna um entreposto para o comércio de escravizados.

O segundo ciclo, a partir da segunda metade do século XIX, coincide com a abolição da escravatura na ilha (1869) e com a abolição formal da condição jurídica dos libertos (1875). Os dois processos levam a uma crise de mão de obra, e a administração colonial passa a adotar o regime de contrato, com a contratação de trabalhadores oriundos de outras colônias portuguesas em África (Angola, Cabo Verde e Moçambique).

Do ponto de vista da caracterização linguística do segundo ciclo de colonização, o período é marcado pela consolidação do Português como L1 dominante, se sobrepondo às línguas crioulas, que possuíam hegemonia absoluta até o século XIX. Gonçalves e Hagemeyer (2015: 88) afirmam que, até o século XVIII, o Forro é a língua materna de grande parte da população nativa de São Tomé, estando o Português em um espaço limitado, como L2. A partir da chegada do grande contingente de contratados na segunda metade do século XIX, o multilinguismo se acentua, mas o contingente populacional que chegou à ilha para o trabalho nas culturas de café e cacau adotou o Português, e não o Forro, como L2.

Durante a colonização, o Português era de acesso muito limitado para os são-tomenses. A partir do Estado Novo em Portugal (1933-1974), a política linguística imposta à colônia foi pautada em uma forte repressão às línguas crioulas, consideradas como ameaça para os interesses do regime. Com o propósito de maior integração à estrutura colonial, a elite urbana são-tomense usava fundamentalmente o português, ainda que se tenham relatos de que os membros dessa elite fossem bilíngues. Contudo, o momento decisivo para a nativização do português é a independência (1975), com sua escolha como língua oficial exclusiva do arquipélago, o que leva à massiva escolarização em português.

Gonçalves e Hagemeyer (2015, p. 91) sintetizam a situação multilinguística atual de São Tomé e Príncipe nos seguintes termos:

De língua da elite e dos domínios altos, o português passou a ser a língua de todos os domínios comunicativos, altos e baixos, da maioria dos são-tomenses. A atual hegemonia do português nas ilhas é também promovida pela ausência de uma política pró-crioula sustentada. A questão da valorização das línguas nacionais veio muitas vezes à tona, mas não produziu estratégias concretas e duradouras para o futuro. Desta forma, a estigmatização dos crioulos, herdada do tempo colonial, não foi devidamente ultrapassada, impedindo, em definitivo, a criação de uma identidade crioula ligada às línguas crioulas. Por todas essas razões, São Tomé e Príncipe é hoje a ex-colônia portuguesa onde

se registra o maior número de falantes nativos do português, o que significa também que todos os crioulos autóctones de São Tomé e Príncipe estão ameaçados.

A situação linguística de Moçambique está configurada de forma bastante distinta da de São Tomé. Moçambique “entra” no cenário colonial português em 1497, com a chegada de Vasco da Gama. Enquanto colônia, Moçambique não despertou, nos séculos XVI e XVII, o interesse do Império Português, tanto que a administração do território era feita a partir da Índia até meados do século XVIII. Uma presença mais efetiva da máquina colonial se registra a partir de 1918, com o fim das campanhas militares para a ocupação efetiva do território. Só a partir de então começam a surgir medidas para a implementação de um sistema de educação.

A partir desse momento, se devolvem as bases para a difusão do Português no território moçambicano. A construção, por parte do governo português, de uma política efetiva de assimilação cultural se estabelece a partir de 1930, com a introdução do Português como língua para instrução escolar. Com a independência, em 1975, o Português é adotado como língua oficial.

Contudo, a adoção do Português como língua oficial se enquadra em um espectro mais amplo da situação multilinguística de Moçambique. Dados do Censo (2007) e de diversos estudos sobre a realidade linguística moçambicana revelam que o Português coexiste com um conjunto muito variado de línguas autóctones, todas pertencentes à família de línguas *Banto*. De acordo com Pissurno (2018, p. 82-83),

Esses idiomas, [...], são, para muitos habitantes das áreas rurais de Moçambique, especialmente aqueles acima dos 50 anos de idade, suas línguas maternas. [...]. Sendo assim, o idioma tido como oficial apresenta um *status* de língua estrangeira (LE) para esses indivíduos, ou seja, uma língua utilizada em situações bastante artificiais, especialmente instrucionais, já que a língua alvo só é aprendida em contextos de educação formal, enquanto em casa os indivíduos utilizam suas línguas maternas para comunicação diária. Por outro lado, nas áreas urbanas, a situação é similar à do uso de uma segunda língua (L2), ou seja, a exposição à língua alvo não se faz apenas em contexto escolar, mas é exigida em praticamente todos os ambientes nos quais os indivíduos estabelecem comunicação, já que, mesmo que dentro de casa eles falem sua língua materna, fora dela é necessário comunicar-se exclusivamente em outra língua, que não sua L1.

Os dados demográficos e descritivos permitem afirmar que a população moçambicana é, em sua maioria, no mínimo bilíngue. A diversidade linguística que caracteriza o território é fruto da política colonial adotada para a região

desde a chegada dos portugueses, que trataram Moçambique como uma área de menor interesse em termos de exploração, e se acentuou com as ações adotadas a partir do início do século XX, com a política de implementação do Português como língua do império.

Com o breve perfil sócio-histórico-linguístico das comunidades investigadas, fica evidente que há pontos de interseção e de separação entre o Português de São Tomé e Príncipe e o Português de Moçambique. Em comum às duas, a emergência de variedades do português em realidades multilíngues. Todavia, as diferenças no processo de exploração colonial levaram ao estabelecimento de políticas diferenciadas para a implementação do Português em cada território. Hagemeyer (2018, p. 8) assim resume as diferenças entre o PST e o PM:

trata-se de duas variedades com características sociolinguísticas distintas: [a variedade do Português de Moçambique] está em contacto com línguas aglutinantes do grupo banto e apresenta uma taxa de nativização crescente, embora ainda relativamente baixa; a [variedade do Português de São Tomé] constitui a L1 da maioria da população, mesmo que os censos nacionais não explicitem esta informação.

## HIPÓTESES E FUNDAMENTAÇÃO TEÓRICO-METODOLÓGICA

Partindo-se da hipótese de que o contato com outras línguas seria um condicionamento essencial na realização de um padrão acentual atípico mesmo para indivíduos que possuem português como L1<sup>3</sup> e que não convivem em situação de multilinguismo generalizado, espera-se que índices elevados de apagamento da vogal postônica medial sejam encontrados tanto no PST quanto no PM.

No âmbito de São Tomé e Príncipe, sabe-se que no Forro a tendência geral é a de apagamento de segmentos, em favor da regularização dos vocábulos a sequências dissilábicas. Sobre essa característica do Forro, vale destacar as considerações de Ferraz (1979, p. 38), o primeiro trabalho de descrição desse crioulo de base lexical portuguesa:

Palavras portuguesas têm tipicamente um número de sílabas maior do que as palavras do Forro. As palavras do Forro são, geralmente, dissilábicas. Essa diferença na estrutura silábica resulta em um **frequente apagamento de segmentos quando palavras portuguesas são incorporadas ao Forro**. Esse apagamento – na forma de aférese,

---

<sup>3</sup> Araújo *et al.* (2007: 37-38) realizaram um levantamento da produtividade dos padrões acentuais, com base no levantamento de todos os verbetes do *Dicionário Houaiss da Língua Portuguesa*. Os autores chegaram a um total de 150.875 palavras, das quais 18.413 são proparoxítonas, o que equivale a somente 12% do total.

síncope ou apócope – pode envolver vogais átonas, sílabas ou seqüências de segmentos. Segmentos acentuados em português não são apagados. (grifo nosso)

Sobre a variedade moçambicana, não se conhecem estudos que descrevam os sistemas vocálicos das línguas banto de Moçambique. Assim, a hipótese de investigação vai na linha de que as proparoxítonas se regularizariam com frequência por conta da baixa frequência de itens lexicais com acento na antepenúltima sílaba e, por consequência, da não naturalidade do padrão proparoxítono no rol dos parâmetros acentuais do Português. É esperado que um indivíduo que tenha Português como L2 ou que tenha adquirido Português como L1 e domine uma (ou várias) língua(s) banto tenha dificuldade em produzir um padrão acentual atípico no inventário fonológico do Português.

Para testagem das hipóteses, levantaram-se 554 dados em 12 inquéritos – seis relativos a São Tomé e seis relativos a Moçambique. Os indivíduos – seis homens e seis mulheres, todos pertencentes à faixa etária mais jovem (18 a 35 anos) – estão distribuídos em três níveis de escolaridade. Nos quadros a seguir, observa-se a distribuição dos informantes, de acordo com a frequência (baixa ou média) de uso de crioulo (São Tomé) e ao uso de Português como L1 ou L2 (Moçambique)

**Quadro 1:** Informantes são-tomenses, distribuídos em relação à frequência de uso de um crioulo

Nível 1 de Instrução		Nível 2 de Instrução		Nível 3 de Instrução	
homem	mulher	homem	mulher	homem	mulher
A1H - baixa	A1M - média	A2H - média	A2M - baixa	A3H - baixa	A3M - baixa

**Quadro 2:** Informantes moçambicanos, distribuídos em relação ao uso do Português como L1 ou L2

Nível 1 de Instrução		Nível 2 de Instrução		Nível 3 de Instrução	
homem	mulher	homem	mulher	homem	mulher
A1H - L2	A1M - L1	A2H - L1	A2M - L2	A3H - L1	A3M - L1

A análise empreendida neste trabalho toma por base o suporte teórico metodológico da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968). Os 554 dados foram analisados com o auxílio do *software* Goldvarb-X. Postularam-se nove variáveis linguísticas – a natureza dos contextos antecedente e subsequente, os traços de articulação das vogais tônica, postônica medial e postônica final, a estrutura da sílaba tônica, a classe morfológica do vocábulo, o

número de sílabas da palavra e a produtividade do item no léxico, e cinco variáveis sociais – sexo, escolaridade, frequência de uso de um crioulo (São Tomé), uso do Português como língua primeira ou segunda (Moçambique) e línguas dominadas pelos informantes (Moçambique).

## RESULTADOS

Os índices gerais e aplicação da regra de apagamento da vogal postônica medial, em função de cada variedade analisada, estão expostos na tabela 1, a seguir.

**Tabela 1:** Índices gerais de apagamento da vogal postônica medial em cada variedade analisada

Variedade	apl/t	exemplo
PST	111/288 = 38%	[ˈsɛklɪ]
PM	96/266 = 36%	[ˈarvriʃ]

A tabela acima revela aspectos convergentes na comparação entre as variedades: não há diferença, em termos de frequência bruta, entre PST e PM nos índices gerais de aplicação da regra. Ambas variedades apresentam índices expressivos de apagamento da vogal postônica medial.

No quadro 3, a seguir, apresentam-se os condicionamentos apontados como estatisticamente relevantes para a aplicação da regra em cada variedade.

**Quadro 3:** Condicionamentos estatisticamente relevantes para o apagamento da postônica medial no Português de São Tomé (PST) e no Português de Moçambique (PM)

PST		PM	
<b>Contexto precedente</b> <i>Escolaridade</i>		<i>Escolaridade</i>	
<u>Contexto subsequente</u>		<b>Contexto precedente</b>	
Frequência de uso de um crioulo		<u>Contexto subsequente</u>	
Input:	.38	Input:	.36
	.30		.35
Significância	.03	Significância	.00

Para a descrição dos resultados, em um primeiro momento serão descritas as variáveis linguísticas selecionadas, um ponto de convergência entre as duas variedades; em um segundo momento, serão descritos os resultados para variável *escolaridade*, também selecionada tanto para o PST quanto para o PM; por fim,

apresentam-se os resultados para a variável *frequência de uso de um crioulo* e uma tentativa de reflexão sobre o papel do *Português como L1 ou L2* no comportamento dos dados moçambicanos.

Na tabela 2, a seguir, apresentam-se os índices percentuais e os pesos relativos para a variável contexto precedente, a variável linguística mais saliente para ambas as variedades.

**Tabela 2:** Consoante precedente

contexto	Exemplos	PST		PM	
		apl/t	P.R	apl/t	P.R
obstruintes	<i>época</i> – [‘ɛpkɐ]	100/196 = 51%	<b>.66</b>	77/165 = 46.7%	<b>.59</b>
nasais	<i>nômade</i> – [‘nõmd]	6/32 = 18.8%	.19	8/43 = 18.6%	.20
líquidas	<i>célula</i> – [‘sɛlɐ]	3/38 = 7.9	.09	6/17 = 35.3%	.42

Os resultados da tabela 2 revelam que, tanto para o PST quanto para o PM, a presença de consoantes obstruintes [ɛpkɐ] no ataque da sílaba postônica não final é o contexto que mais favorece o apagamento da vogal postônica medial, com .66 e .59 de peso relativo, respectivamente. Os demais contextos considerados, nas duas variedades, estão abaixo do ponto neutro, se mostram desfavorecedores para a ocorrência do processo. Quando se contrastam os resultados do contexto precedente com o contexto subsequente, percebe-se que há uma certa convergência entre as variáveis. Os resultados para as consoantes subsequentes estão expressos na tabela 3, a seguir.

**Tabela 3:** Consoante subsequente

contexto	Exemplos	PST		PM	
		apl/t	P.R	apl/t	P.R
obstruintes	<i>político</i> – [p’litku]	79/210 = 37.6%	.45	54/170 = 31.8%	.46
nasais	<i>décimo</i> – [‘desmu]	22/53 = 41.5%	.54	26/68 = 38.2%	.43
líquidas	<i>espetáculo</i> – [ʃp’taklu]	10/25 = 40%	<b>.77</b>	16/28 = 57.1%	<b>.82</b>

Os índices expressos na tabela 3 mostram que as consoantes líquidas – a lateral e a vibrante – são os contextos que mais favorecem o apagamento da vogal postônica medial nas duas variedades (.77 para o PST e .82 para o PM). Nos dados do PST, ainda há um ligeiro favorecimento proporcionado pela presença de uma consoante nasal no ataque da sílaba átona final (.54). Quando se compararam tais índices com os verificados para o contexto precedente – variável em que

as obstruintes se revelam como um contexto favorecedor de aplicação da regra nas duas variedades – nota-se que o apagamento é regido pela possibilidade de ressilabificação das consoantes em direção ao ataque da sílaba átona final, com a formação de um ataque complexo formado por obstruinte e consoante líquida. Essa é uma tendência histórica em português, com raízes no latim vulgar<sup>4</sup>.

Com relação à variável escolaridade, selecionada tanto para o PM quanto para o PST, os resultados já não se revelam uniformes para as duas variedades. Na tabela 4, a seguir, é possível perceber que o nível de instrução é um condicionamento com efeito diferenciado para cada conjunto de dados.

**Tabela 4:** Escolaridade

	PST		PM	
	apl/t	P.R	apl/t	P.R
Nível 1	7/67 = 10.4%	.13	12/50 = 24%	.39
Nível 2	43/94 = 45.7%	.53	64/116 = 55.2%	<b>.70</b>
Nível 3	61/129 = 47.3%	<b>.70</b>	20/100 = 20%	.30

Para os dados de São Tomé, percebe-se uma escalaridade na atuação da variável: quanto maior o nível de instrução, maiores são os pesos relativos para o apagamento. A expectativa era que o aumento do nível de instrução atuasse como um bloqueador da regra de apagamento. Não podemos nos esquivar da hipótese de que o maior favorecimento para o nível 3 de instrução tenha alguma relação com uma espécie de espelhamento da norma de referência do Português de São Tomé, que seria a norma do Português Europeu, um sistema que propicia o apagamento de segmentos átonos. Os indivíduos com os níveis mais altos de instrução são aqueles que saíram de São Tomé e foram estudar em Portugal. Contudo, o Forro também se caracteriza por um processo amplo de apagamento de segmentos átonos para a implementação do padrão dissilábico. Logo, a questão da atuação da variável escolaridade pode estar atrelada a questões outras, muito complexas para serem medidas a partir de variantes tão estanques.

<sup>4</sup> Destaca-se que o Appendix Probi (CASTRO, 1981, p. 81-83) já registra formas do tipo *speculum non speclum, masculus non masclus, vetulus non veclus, vitulus non viclus, vernaculus non vernachus, articulus non articlus, baculus non vaclus, angulus non anglus, iugulus non iuglus, barbarus non barbar, calida non calda, oculus non oclus, stabulum non stablum, tribula non tribla, viridis non virdis, vapulo non baplo*. Na maior parte dos dados, as formas “condenadas” sofrem redução pela queda da vogal postônica medial e pela formação de uma estrutura consoante obstruinte + líquida no ataque da sílaba átona final.

Os resultados para os dados de Moçambique poderiam ser ainda mais contraditórios, caso tratados de forma isolada. Um olhar atento para o perfil dos informantes com nível médio de instrução pode justificar a relevância da variante e a seleção da variável na análise estatística.

Os dois informantes de nível médio são os que produzem o maior número de *tokens* (116) e o maior número de palavras proparoxítonas diferentes – 27 itens lexicais para o informante masculino e 29 para a informante feminina. Além disso, cumpre destacar que os dois informantes, pela profissão que exercem (são funcionários de um hotel), possuem uma desenvoltura maior do que os demais que compõem a subamostra deste trabalho. Os dois informantes são mais espontâneos, se arriscam mais em termos de produtividade de palavras proparoxítonas e de quantidade de dados. Por consequência, exibem um índice maior de apagamento da vogal átona medial, o que se demonstrou de forma contundente na análise estatística. Logo, em Moçambique, o efeito não é exclusivo da escolaridade, mas um reflexo do perfil social dos informantes.

A influência do contato linguístico no apagamento da vogal postônica medial foi controlada através de três variáveis: a frequência de uso de um crioulo (São Tomé), o uso de Português como L1 ou L2 e as línguas dominadas pelos informantes (Moçambique). Das três variáveis postuladas, apenas a relativa aos dados são-tomenses foi selecionada pelo programa de análises estatísticas. Na tabela 5, a seguir, apresentam-se os índices percentuais e os pesos relativos para a frequência de uso de um crioulo.

**Tabela 5:** Frequência de uso de um crioulo<sup>5</sup> (PST)

	apl/t	P.R
Frequência baixa	78/207 = 37.7%	.43
Frequência média	33/50 = 66%	<b>.66</b>

Na tabela acima, é possível observar que os indivíduos que se comunicam fundamentalmente em Português são aqueles que bloqueiam a aplicação da regra de apagamento da vogal postônica medial (.43), enquanto os falantes que afirmam usar o Forro eventualmente são os que favorecem a implementação do

<sup>5</sup> A variável Frequência de uso de um crioulo foi formulada por Brandão (2016, p. 91) nos seguintes termos: “frequência (a) zero/baixa, referente aos indivíduos que se expressam fundamentalmente em Português; (b) média, relativa aos indivíduos que se expressam em Português, mas dominam um crioulo e dele fazem uso eventualmente; (c) alta, abarcando os indivíduos que, embora falem o Português e o tenham como L1, se expressam, regularmente, num crioulo”.

apagamento (.66). De certa forma, os resultados se coadunam com a hipótese inicial, da influência de uma tendência do Forro ao apagamento de segmentos átonos nas palavras de origem portuguesa com mais de duas sílabas, processo descrito por Ferraz (1979). Contudo, com base nas informações do quadro 1, nota-se que apenas dois informantes se declaram como usuários eventuais do Forro (o homem do nível 1 e a mulher do nível 2 de instrução). Somente com a ampliação do *corpus*, os efeitos da variável ficam mais evidentes<sup>6</sup>.

Com relação aos dados do PM, as variáveis que controlam os efeitos do contato linguístico não foram selecionadas. Os resultados estão expressos na tabela 6, a seguir, e podem fornecer diretrizes para justificar a não seleção das variáveis.

**Tabela 6:** Relação entre o Português e as línguas locais<sup>7</sup>

Uso de Português como L1 ou L2		
	apl/t	P.R
Português como L1	61/180 = 33.9%	(.53)
Português como L2	35/80 = 40.7%	(.42)
Línguas dominadas pelos informantes <sup>8</sup>		
Só Português ou apenas compreende as línguas locais	39/93 = 41%	(.46)
Fala Português e línguas locais	57/173 = 32.9%	(.52)

A distorção entre os índices percentuais e os pesos relativos pode justificar a não seleção estatística das variáveis que controlam os efeitos do contato linguístico nos dados do PM. Nota-se que, em termos percentuais, a variante que dá conta do uso de Português como L2 é a mais produtiva no conjunto de dados (40.7%), mas a que desfavorece, em termos de peso relativo, a aplicação da regra de apagamento (.42); o mesmo processo se observa para a variável que investiga as línguas usadas pelos informantes, uma vez que a variante que engloba os dados dos indivíduos que se reconhecem como falantes de Português e línguas

<sup>6</sup> Gomes (2018) faz uma descrição do efeito da variável frequência de uso de crioulos em um conjunto maior de dados, que incluem os indivíduos da segunda faixa etária. Os resultados, no referido estudo, convergem com os índices preliminares apresentados neste trabalho.

<sup>7</sup> Os pesos relativos para as variáveis descritas na tabela foram extraídos da primeira rodada do step-down.

<sup>8</sup> Variável formulada originalmente no trabalho de Pissurno (2017), sobre a concordância verbal nos dados do mesmo corpus.

locais é a menos frequente (32.9%), e a que favorece o apagamento da postônica medial em peso relativo (.52). A falta de convergência entre frequência e peso relativo pode ser um reflexo da atuação de outros parâmetros, que impediram a seleção pelo programa de análises estatísticas.

## PARA ENCERRAR (OU ... PARA INICIAR A CONVERSA ...)

Neste trabalho, formulado com o intuito de explorar preliminarmente duas realidades linguísticas dotadas de alta complexidade, buscou-se observar se haveria convergências entre as variedades são-tomense e moçambicana do Português na atuação da regra de apagamento da vogal postônica medial, e se seria possível conduzir um trabalho comparativo entre essas duas variedades, que se caracterizam por conviverem em contexto de multilinguismo acentuado (mesmo que a sócio-história de cada comunidade revele traços diferenciados na relação entre o Português e as línguas locais).

Em termos de frequência bruta de aplicação da regra, nota-se um comportamento similar entre as variedades: índices expressivos de apagamento da vogal átona medial (38% para o PST e 36% para o PM). Observaram-se similaridades na atuação das variáveis linguísticas: em ambas, os contextos adjacentes à vogal postônica são os condicionamentos decisivos para a implementação do apagamento, uma vez que a possibilidade de ressilabificação das consoantes precedente e subsequente em direção ao ataque da sílaba átona final é a restrição que condiciona o apagamento nos dois conjuntos de dados.

Divergências se verificam na atuação das variáveis sociais. A escolaridade atua de forma diferenciada para cada variedade: enquanto nos dados de São Tomé há um movimento escalar, no sentido em que os índices de apagamento aumentam com o aumento da escolaridade, os dados de Moçambique mostram um favorecimento da regra pelos indivíduos de nível intermediário de instrução. Todavia, os dados moçambicanos se revelaram muito mais sensíveis ao perfil social dos informantes do nível médio do que propriamente um efeito do acesso à norma europeia.

Outra diferença se verificou na atuação das variáveis que controlam o papel do contato linguístico no apagamento da vogal átona medial: os dados são-tomenses se mostraram sensíveis ao efeito da frequência de uso de um crioulo. Em Moçambique, por outro lado, as variáveis que davam conta da relação do Português com as línguas locais não se mostraram relevantes, até por conta da falta de convergência entre a frequência bruta e os pesos relativos.

Em síntese, os resultados deste estudo inicial revelam que é possível o tratamento comparativo entre as variedades do Português faladas em São Tomé e em Moçambique na análise do fenômeno variável em foco. Entretanto, os resultados aqui discutidos jamais podem ser tomados como absolutos. É fundamental a adoção de procedimentos teórico-metodológicos que ultrapassem uma análise meramente quantitativa, e que se direcionem para uma análise mais verticalizada de cada comunidade. Essa investigação mais complexa está em andamento. A lição que este trabalho deixa é a de que muito há a se fazer na investigação sobre aspectos fonético-fonológicos das variedades africanas do Português, mas o que já se descreveu sobre São Tomé e Moçambique deixa evidente a necessidade de atrelar a descrição linguística à realidade multilíngue dessas comunidades.

## REFERÊNCIAS

AMARAL, Marisa Porto do. *As proparoxítonas: teoria e variação*. Porto Alegre, PUC-RS, 2000. 235.fl.s. Tese (Doutorado em Letras). Pontifícia Universidade Católica do Rio Grande do Sul.

ARAÚJO, Gabriel Antunes de *et. al.* As proparoxítonas e o sistema acentual do português. In: ARAÚJO, Gabriel Antunes de *et al.* (org.). *O acento em português: abordagens fonológicas*. São Paulo: Parábola, 2007. p. 37-60.

BRANDÃO, Sílvia Figueiredo. Variação e o estatuto de variedades do Português. *Diadorim*. v.18, Número Especial, p. 83-104, 2016.

CAIXETA, Valmir. *Descrição e análise da redução das palavras proparoxítonas*. 1989. Dissertação (Mestrado em Linguística). Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1989.

CASTRO, Ivo. *Curso de história da língua portuguesa*. Lisboa: Universidade Aberta, 1991.

CHAVES. Raquel Gomes. *A redução de proparoxítonos na fala do Sul do Brasil*. 2011. 173f. Dissertação (Mestrado em Linguística). Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2011.

FERNANDES, Ana Catarina Garcia. *Apagamento de vogais átonas em trissílabos proparoxítonos: um contributo para a compreensão da supressão vocálica em português europeu*. 2007. 155 f. Dissertação (Mestrado em Linguística). Universidade do Porto, Porto, 2007.

FERRAZ, Luiz Ivens. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.

FONSECA, Simone Meckler. *O problema das proparoxítonas: a perda da vogal postônica*. 2007. 68 f. Dissertação (Mestrado em Linguística). Universidade Federal de Juiz de Fora, Juiz de Fora, 2007.

GOMES, Danielle Kely. *Síncope em proparoxítonas: um estudo contrastivo entre o português brasileiro e o português europeu*. Rio de Janeiro: UFRJ, 2012. 273 f. Tese de Doutorado em Letras Vernáculas. Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2012.

GOMES, Danielle Kely. O apagamento das vogais postônicas não finais: uma análise contrastiva entre variedades do português. *Revista da ABRALIN*, vol. 14: 185-106, 2015.

GOMES, Danielle Kely. Síncope das vogais postônicas não finais: uma análise contrastiva entre variedades do português. In: DE PAULA, Alessandra; GOMES, Danielle Kely; SILVEIRA, Eliete; MACHADO VIEIRA, Marcia dos Santos; VIEIRA, Silvia Rodrigues. (orgs). *Uma história de investigação sobre a Língua Portuguesa: homenagem a Silvia Brandão*. São Paulo, Blucher, 2018. p. 213-224.

GOMES, Danielle Kely. Vogais em contexto postônico medial no português de São Tomé. In: BRANDÃO, Silvia Figueiredo. (org). *Dois variedades africanas do português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo, Blucher, 2018. p. 159-176.

GONÇALVES, Perpétua. *Português de Moçambique: uma variedade em formação*. Maputo: Livraria Universitária da Universidade Eduardo Mondlane, 1996.

GONÇALVES, Rita; HAGEMEIJER, Tjerk. O Português num contexto multilíngue: O caso de São Tomé e Príncipe. *Revista Científica da Universidade Eduardo Mondlane*, Moçambique, v.1, n. 1, p. 87-107, 2015.

HAGEMEIJER, Tjerk. As ilhas de Babel: A criouliização no Golfo da Guiné. *Revista Camões*. 6, p. 74-88, 1999.

HAGEMEIJER, Tjerk. As línguas de São Tomé e Príncipe. *Revista de Crioulos de Base Lexical Portuguesa e Espanhola* (1)1, p. 1-27, 2009.

HAGEMEIJER, Tjerk. Prefácio. In: BRANDÃO, Silvia Figueiredo. (org). *Dois variedades africanas do português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher, 2018, p. 7-9.

LIMA, Giselly de Oliveira. *O efeito da síncope em proparoxítonas: uma análise fonológica e variacionista com dados do sudoeste goiano*. 2008. 195 f. Dissertação (Mestrado em Linguística). Universidade Federal de Uberlândia, Uberlândia, 2008.

LIMA, Giselly de Oliveira. *Percepção da síncope em palavras proparoxítonas*. 2017. 154f. Tese (Doutorado em Linguística). Universidade Federal de Uberlândia, Uberlândia, 2017.

NASCIMENTO, Fabiane de Mello Vianna da Rocha Teixeira Rodrigues do. *O sistema vocálico do português de São Tomé e o comportamento das vogais médias pretônicas*. 2018. 194 f. Tese (Doutorado em Língua Portuguesa). Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2018.

PISSURNO, Karen Christina da Silva. *A concordância verbal de terceira pessoa do plural na variedade moçambicana do Português: uma abordagem sociolinguística*. 2017. 213 fls. Dissertação (Mestrado em Língua Portuguesa). Rio de Janeiro: Faculdade de Letras da UFRJ, 2017.

PISSURNO, Karen Christina da Silva. O perfil multilíngue de Moçambique. In: BRANDÃO, Silvia Figueiredo. (org). *Dois variedades africanas do português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher. p. 75-91.

RAMOS, Adriana Perpétua. *Descrição das vogais postônicas não finais na variedade do Noroeste Paulista*. 2009. 175f. Dissertação (Mestrado em Linguística). Universidade Estadual Paulista, São José do Rio Preto, 2009.

SILVA, André Pedro. *Supressão da vogal átona postônica não-final: uma tendência das proparoxítonas na língua portuguesa com evidências no falar sapeense*. 2006. 139 f. Dissertação (Mestrado em Língua Portuguesa). Universidade Federal da Paraíba, João Pessoa, 2006.

SILVA, André Pedro. *Vogais postônicas não finais: do sistema ao uso*. João Pessoa, 2010, 216f. Tese (Doutorado em Linguística). Universidade Federal da Paraíba, João Pessoa, 2010.

TIMBANE, Alexandre António. Que português se fala em Moçambique? Uma análise sociolinguística da variedade em uso. In: *Revista Vocábulo*, v. VII. São Paulo, 2014.

WEIREICH, Uriel; LABOV, William; HERZOG, Marvin. Empirical foundations for a theory of language change. In: LEHMANN, W. MALKIEL, Y. (Eds.). *Directions for historical linguistics*. Austin: University of Texas Press, 1968. p. 97-195.

## THE DELETION OF POSTONIC MEDIAL VOWEL IN TWO AFRICAN VARIETIES OF PORTUGUESE A PRELIMINARY STUDY

*Danielle Kely Gomes (UFRJ)*

The deletion of the medial postonic vowel, a process that culminates the regularization of proparoxytone words to the paroxytone pattern, is an old phenomenon in Portuguese. Gomes (2012), in a variationist analysis that contrasts the Brazilian Portuguese (BP) with the European Portuguese (EP), finds that in the EP erasure rule is more frequent than in the BP, a variety that can be considered conservative in terms of the configuration of the systems vowels, given the variation that still occurs in the context of the realization of the medium vowels in the unstressed pretonic context and the negative social valuation to which vowel deletion is subject. Gomes (no prelo), in a comparison between BP, EP and São Tomé Portuguese (STP), identifies in this a very particular behavior: high vowel deletion index in the postonic medial context, probably a reflection of the influence of creole that coexists with Portuguese in the area of data collection.

This paper proposes a comparison between the informants of the first São Tomé age group with data from individuals of the same profile extracted from the Portuguese Mozambican corpus (collected in Maputo in 2016). In Mozambique, Portuguese coexists with a wide variety of languages of the Bantu family. The hypothesis is that proparoxytones, unnatural even for speakers of Portuguese as L1 and that are not immersed in multilingual contexts, would be regularized to

paroxitones in the Mozambican variety - as is verified for the São Toméan data - as a contact effect of Portuguese with the other languages that coexist with it in the community.

For the hypothesis testing, 554 data were collected in 12 recorded data - six in São Tomé and six in Mozambique. The sub-sample for this preliminary study were stratified according to the variables gender and schooling. Data were also categorized according to the frequency of use of local languages (in São Tomé) and the acquisition of Portuguese as the first or second language (in Mozambique). The data, analyzed from the theoretical-methodological foundations of Theory of Variation and Change, were treated statistically from the GOLDVAR-BX program package. In this work, the effects of nine structural variables and five social variables were investigated.

The results reveal, in general terms, a similar behavior among the varieties: in both São Tomé and Mozambique there were expressive indices of erasure of the medial atonic vowel (around 30%). Similarities were observed in the performance of linguistic constraints. In both varieties, the contexts adjacent to the medial postonic vowel were decisive for the implementation of erasure, since the possibility of restructuration of the preceding and subsequent consonants towards the attack of the final atonic syllable is the constraint that conditions the deletion in the two sets of data.

Divergences were observed in the performance of social variables. Schooling acts differently for each variety. In the São Tomé data, the erasure rates of the medial postonics increase with the increase in schooling. The Mozambican data, on the other hand, reveal a favoring of the rule by the individuals of intermediate level of instruction. However, the results verified for the Mozambican speakers reflect much more the communicative profile of the informants than an effect of the contact with what would be prestigious in terms of norm.

Another difference is consistent in the performance of the variables that control the role of linguistic contact in the erasure of the medial atonic vowel. In São Tomé, contact with local languages seems to favor the erasure of the medial postonics. In Mozambique, however, the variables that account for the relationship between Portuguese and local languages were not relevant.

The results of this preliminary study show the possibility and necessity of a comparative treatment between the varieties of Portuguese spoken in São Tomé and in Mozambique in the analysis of the variable phenomenon of deletion of the medial postonic vowel. However, the results discussed here should not be taken as conclusive. It is recognized that there is a need for methodological procedures

that go beyond a purely quantitative description and that direct the study to a vertical analysis of each variety.

## REFERENCE

GOMES, Danielle Kely. *Síncope em proparoxítonas: um estudo contrastivo entre o português brasileiro e o português europeu*. Rio de Janeiro: UFRJ, 2012. 273 f. Tese de Doutorado em Letras Vernáculas. Faculdade de Letras, Universidade Federal do Rio de Janeiro, 2012.

GOMES, Danielle Kely. O apagamento das vogais postônicas mediais em continuum: uma comparação entre as variedades brasileira, europeia e são-tomense. In: BRESCANCINI, Claudia Regina; MONARETTO, Valéria. E-book do Encontro Intermediário do GT de Sociolinguística da ANPOLL. (no prelo).



## AINDA SOBRE OS RÓTICOS NO PORTUGUÊS DE MOÇAMBIQUE

*Silvia Figueiredo Brandão (UFRJ/CNPq)*  
*Alessandra de Paula (UERJ/FAPERJ)*

### RESUMO

Focalizam-se os róticos no âmbito da variedade urbana do Português de Moçambique (PM) à luz dos pressupostos teórico-metodológicos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968). Objetiva-se discutir os fatores estruturais e sociais que condicionam o uso das variantes de R na posição pré-vocálica – como em *rosa* e *carro* – na fala de indivíduos que o utilizam como L1 e L2. Realizaram-se análises controlando variáveis estruturais e também sociais, como sexo, faixa etária/nível de escolaridade e as línguas faladas pelos informantes, esta última com o propósito de aquilatar as possíveis interferências das línguas locais, devido à situação multilinguística da área. Os resultados das análises demonstram que o tepe, a variante dominante no PM, ocorre na fala de todos os indivíduos, mas é significativamente mais frequente entre aqueles que o têm como L2. Entre os falantes de PM L1, sobretudo os de nível superior de escolaridade, prevalece a vibrante alveolar, o que os aproxima da norma europeia. Apesar dessas constatações, acredita-se que na gramática do PM ocorra um único fonema rótico.

## INTRODUÇÃO

No Português de Moçambique (PM), a distribuição dos róticos não equivale à que se observa no Português do Brasil (PB) e no Português Europeu (PE), aparentemente perdendo-se, em contexto intervocálico, a possibilidade de oposição significativa (*caro x carro*), quadro que se supõe poder derivar-se do acentuado multilinguismo<sup>1</sup>, e consequente complexidade sociolinguística, que caracteriza a sociedade moçambicana. As primeiras análises, já divulgadas por Brandão; De Paula (2017; 2018), demonstram que nos contextos pré-vocálicos (*roça, carro*) e de coda externa (*falar, flor*) predomina o tepe, secundado pela vibrante alveolar, no primeiro caso, e pelo apagamento, no segundo, levando em conta todas as amostras.

Neste estudo, dá-se continuidade às análises efetuadas, desta vez observando-se, em especial, o comportamento dos falantes de Português como L1 (PL1) em contraposição aos que o têm como L2 (PL2). Os resultados das primeiras análises motivaram a presente investigação, visto que demonstraram a relevância da variável *Estatuto do Português* (L1 ou L2), especialmente na posição intervocálica. Nesse contexto, o tepe foi a variante mais usada pelos falantes de PL2. Focaliza-se, então, o chamado “R forte” nos contextos inicial de vocábulo e intervocálico, buscando, entre outros objetivos, verificar se são os mesmos os fatores que concorrem, em cada caso, para a implementação do tepe e se outras variantes (especialmente a vibrante alveolar) estariam presentes na gramática dos indivíduos de nível superior desses dois segmentos sociais.

Para desenvolver o tema, além desta introdução, na primeira seção, comenta-se a distribuição dos róticos no PE e nos estudos mencionados na introdução e, na segunda seção, traça-se um breve perfil da área de pesquisa, ressaltando aspectos histórico-sócio-culturais que estão na base da complexidade observada, com apoio, sobretudo, em Chimbutane (2018). Na terceira seção, indicam-se os aspectos teórico-metodológicos que nortearam as análises, bem como os índices do tepe, a variante majoritária, por informante. A quarta seção é dedicada à apresentação dos resultados das análises e a quinta, à sua discussão. Na sexta seção, apresentam-se as considerações finais.

---

<sup>1</sup> A hipótese leva em conta a diferenciada distribuição das diversas línguas do grupo Banto pelo território moçambicano. Lamentavelmente, só se pôde ter acesso a uma dessas gramáticas, referente ao Changana (NGUNGA; SIMBINI, 2012), que predomina em Maputo, a área de recolha de dados.

## OS RÓTICOS

Os róticos têm sido, não apenas no âmbito do Português, alvo de inúmeros estudos, na perspectiva quer fonética, quer fonológica, em virtude não só de sua significativa presença nas línguas do mundo (em cerca de 75% delas), mas também pelo fato de, a despeito de apresentarem tendências similares quanto à posição no vocábulo, não constituírem uma classe no sentido estrito do termo. Na realidade, esse conjunto de sons não apresenta um traço comum a todos eles, sendo o único elemento a caracterizá-los o fato de serem representados pelas letras *r* ou *rr* nas línguas em que ocorrem, o que levou Lindau (1985) a denominá-los de “família de sons”.

Como observa Brandão (2018, p. 391)<sup>2</sup>:

Ladefoged; Maddieson (1996, p. 244-245) mencionam o terceiro formante mais abaixado como um possível traço unificador de todos os róticos, mas acabam por concluir que, na realidade, isso não se aplica, pois não só o abaixamento não ocorre em todos os róticos, mas também “cada membro da classe se assemelha a algum outro membro em relação a alguma propriedade, mas não é a mesma propriedade que constitui a semelhança entre todos os membros da classe”. Para os autores, tais semelhanças parciais poderiam explicar “diversas alternâncias sincrônicas e as mudanças diacrônicas que conectam diferentes tipos de róticos entre si”, acrescentando que “embora haja vários subtipos bem definidos de sons (vibrantes, flapes, etc.) incluídos na classe rótica, a unidade global do grupo parece residir sobretudo nas conexões históricas entre esses subgrupos, e na escolha da letra ‘r’ para representar todos eles”

Embora a complexidade dessa família de sons possa ser exemplificada pelo caso do Português do Brasil, em que se registram fricativas, vibrantes, aproximantes e tepes, com maior variabilidade nos contextos pós-vocálicos, cabe verificar como se comportam no PE, que, a princípio, constitui a norma de referência do PM.

Mateus; d’Andrade (2000, p. 11) indicam para o contexto de oposição fonológica a ocorrência do tepe (como em *caro*) e da vibrante uvular (como em *carro*) na fala atual de Lisboa, a mesma que se encontra em início de vocábulo (*roça*). Segundo os autores, este último segmento coocorre com outras variantes recuadas, “sobretudo a fricativa uvular sonora [ʁ] ou a surda [χ], sendo que a vibrante alveolar [r] é comum em outros dialetos que não o aqui em consideração” (e que seria a variante conservadora).

---

<sup>2</sup> A tradução deste e dos demais trechos citados neste estudo são da responsabilidade das autoras.

Veloso (2015, p. 328), em consonância com esses autores, afirma haver, no PE, um processo de posteriorização do R nos contextos pré-vocálicos, que redundou, na norma lisboeta atual, na inovadora vibrante uvular [R] que estaria deixando de ser uma sonorante, sofrendo “uma mudança subsequente mais drástica”, isto é, estaria sendo substituída por uma fricativa, “dentro de uma gama de escolhas que incluiria velares (desvozeada [x] ou vozeada [ɣ]) e uvulares (desvozeada [χ] e vozeada [ʁ])”. Além disso, ele menciona a emergência, no PE, da aproximante retroflexa, em contexto pós-vocálico, sobretudo na fala de indivíduos de alto grau de escolaridade, na área do Porto.

Veloso finaliza suas observações (p. 334), que também contemplam o PB, apresentando uma síntese da cronologia da mudança dos róticos desde as primeiras constatações de Gonçalves Viana (1883, 1903), e que se reproduz com pequenas alterações formais no Quadro 1.

**Quadro 1:** Evolução dos róticos no PE e no PB, segundo Veloso (2015)

VIBRANTES	Pré e início do século XIX	Vibrante alveolar /r/
	Séculos XIX e XX	Vibrante uvular /R/
	Século XX até o estágio atual	PE: Fricativas [ʁ] > [χ] > [x] (RENNICKE; MARTINS, 2013; escala de frequência) PB: Fricativas e glotais [x ɣ h fi] (SILVA,2002)
FLAPES	Século XIX e início do XX	Flap alveolar [r ]
	A partir de meados do século XX	PE: [r] Emergência do [ɽ] em determinados dialetos e contextos prosódicos PB: [r] Realização muito comum de [ɽ] num crescente número de posições prosódicas e contextos sociais/regionais

Fonte: Veloso (2015: 334), com pequenas alterações formais

Veloso observa, ainda, que [r], [ɽ] e [R] não desapareceram completamente do PE e do PB modernos e que, no quadro, apenas “os alofones inovadores são considerados na linha do tempo de acordo com a suposta data de seu surgimento na língua”.

Quanto às variedades africanas do Português, como destacam alguns de seus estudiosos (HAGEMEIJER, 2018; GONÇALVES, 2010, 2013; CHIMBUTANE, 2018), o plano fonético/fonológico é o menos contemplado nas análises. No âmbito do Português de São Tomé (PST), a partir de 2015, começou a haver interesse pelos róticos (talvez a variável fônica mais saliente em relação ao PE e ao PB), uma vez que se verifica, de um lado, uma grande heterogeneidade, com predomínio do tepe (BRANDÃO et al., 2017; BRANDÃO; DE PAULA, 2018), de outro, sobretudo na fala dos mais jovens, indícios de mudança com a implementação da fricativa uvular sonora [ʁ] (BOUCHARD, 2017; PEREIRA; HAGEMEIJER; FREITAS, 2018). Assim, no PST, coexistem formas como [r]oça/[ʁ]oça/[ʁ]oça com formas como p[r]ofesso[r]a/p[ʁ]ofesso[ʁ]a.

Sobre os róticos no PM só se dispõe de informações esparsas (GONÇALVES, 2013) e dos estudos realizados por Brandão (2018) e Brandão; De Paula (2017; 2018). Nessa variedade, no cômputo geral, o tepe é a variante mais difundida em todos os contextos, seguida pela vibrante alveolar nos chamados contextos de “R forte”.

Brandão; De Paula (2018) analisaram os róticos, nos contextos pré e pós-vocálico, com os mesmos informantes moçambicanos que contribuíram para a presente análise. No entanto, focalizaram-nos numa diferente perspectiva, isto é, considerando os falantes de PL1 e PL2 em conjunto, embora levassem em conta a variável *Estatuto do Português* (L1, L2), cujos resultados acabaram por motivar a consecução deste estudo. Nos Quadros 2 e 3, sintetizam-se os resultados obtidos para a implementação do tepe em posição pré-vocálica.

**Quadro 2:** Síntese de resultados da análise de R forte em contexto inicial de vocábulo no PM

**Quadro 3:** Síntese de resultados da análise de R forte em contexto intervocálico no PM

<p><b>R INICIAL DE VOCÁBULO – PM</b>  <b>Valor de aplicação: Tepe (x vibrante alveolar)</b>  <b>416/752 dados</b>  <b>Tepe: 55,3%</b>  <b>VARIÁVEIS SELECIONADAS:</b>  <i>Nível de escolaridade, Tonicidade da sílaba/ Faixa etária</i></p> <p>Input: .56                      Significância: .001</p>	<p><b>R INTERVOCÁLICO – PM</b>  <b>Valor de aplicação: Tepe (x vibrante alveolar)</b>  <b>254/521 dados</b>  <b>Tepe: 48,8%</b>  <b>VARIÁVEIS SELECIONADAS:</b>  <i>Nível de escolaridade/ Estatuto do Português/ Faixa etária</i></p> <p>Input: .48                      Significância: .000</p>
--	---

Fonte: Brandão; De Paula (2018)

Dos dados expostos, cabe salientar que o tepe no contexto inicial de vocábulo (como em roça) apresenta índice levemente superior ao do contexto intervocálico (como em carro), com apenas uma diferença de 6,5%, e que as variáveis mais representativas para a sua implementação são basicamente de cunho social, conforme a hipótese inicial. Além disso, os resultados parecem indicar que, a exemplo do que ocorre no PST, na gramática da maior parte dos falantes do PM, os róticos [+ant] e [-ant], que atuam no contexto de oposição fonológica no PE e no PB, sofreram um processo de neutralização, o que será discutido na quinta seção.

## A ÁREA DA PESQUISA

Moçambique, situado no Sudeste africano, com 801.537km<sup>2</sup> e 28.861.863 habitantes<sup>3</sup>, está dividido em 11 províncias, entre as quais a da cidade de Maputo, a capital da República, onde se realizaram as entrevistas que servem de base a este estudo. Apesar de, no país, serem faladas mais de 20 línguas do grupo Banto, o Português é a única língua oficial, *status* que adquiriu após a independência em 1975. A essas línguas somam-se o Inglês, o Árabe, o Hindi, o Gujarati e o Urdu (Chimbutane, 2018), o que ainda mais contribui para o multilinguismo e o multiculturalismo que caracterizam a área.

Chimbutane (2018) procura traçar, com base numa perspectiva sócio-histórica, o perfil sociolinguístico de Moçambique, na tentativa de compreender “a gênese e a diversidade do Português Moçambicano” (p. 89-90).

Segundo ele, os portugueses chegaram a Moçambique em 1498 e, mais interessados no comércio com a Índia, passaram a ocupar, a partir de 1505, os centros comerciais estratégicos, sendo o território administrado de Goa, como parte da Índia Portuguesa, até 1752. Ele observa que “embora o que constitui o atual Moçambique date de fins do século XV, só em finais do XIX as atuais fronteiras foram definidas e a relação entre os dois países foi institucionalizada”, sendo que “os portugueses só alcançaram a pacificação e o efetivo controle do território em inícios do século XX” (p. 91). Depois da independência, houve uma série de conflitos, inclusive uma guerra civil, situação que começou a mudar em 1990, com uma nova Constituição, e, mais especificamente, em 1994, com a realização das primeiras eleições democráticas.

No que se refere ao PM, Chimbutane (2018, p. 92) observa que

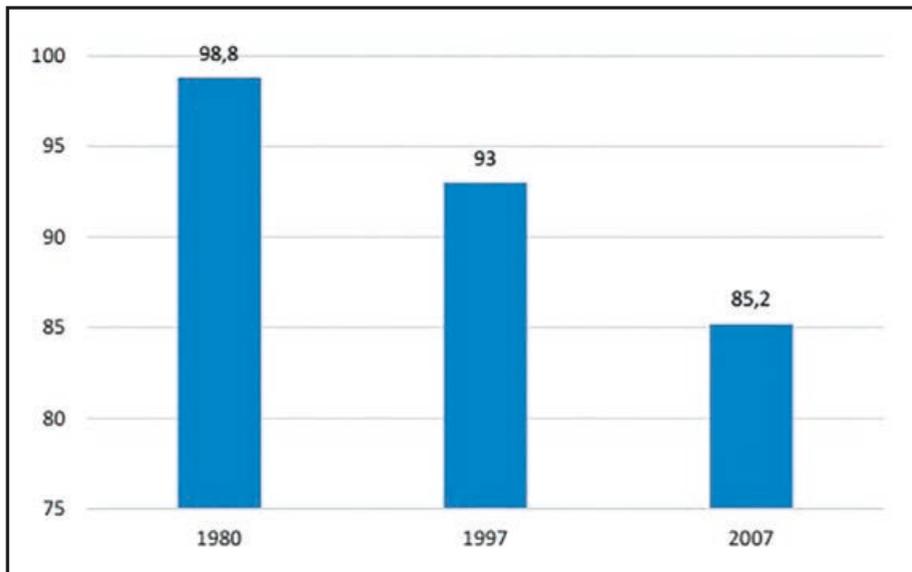
---

<sup>3</sup> Censo de 2017, ainda em processamento. Cf. <http://www.ine.gov.mz/>

Apesar de algumas melhorias após a independência, a guerra civil (1976-1992) e as incertezas políticas do pós-guerra limitaram ainda mais o desenvolvimento da educação em Moçambique. No tempo colonial, o atraso na efetiva colonização reduziu a proporção de nativos que se beneficiaram da educação colonial em Moçambique, com a conseqüente expansão limitada da língua portuguesa. Além disso, movimentos populacionais pós-independência relacionados à guerra, em particular de áreas rurais a urbanas e de uma área etnolinguisticamente ligada a outra, podem ter afetado o curso do desenvolvimento de variedades subnacionais de português, especialmente quando se leva em consideração o papel catalisador das línguas africanas na nativização dessa língua. [...] No caso de Moçambique, isto incluiu o desenvolvimento de novas formas gramaticais e discursivas, muitas vezes influenciadas pelas características das línguas africanas e dos valores socioculturais a elas associados. Embora este processo tenha sido principalmente espontâneo, ele também foi influenciado por discursos ideológicos, incluindo aqueles que fundamentam o projeto de construção da nação.

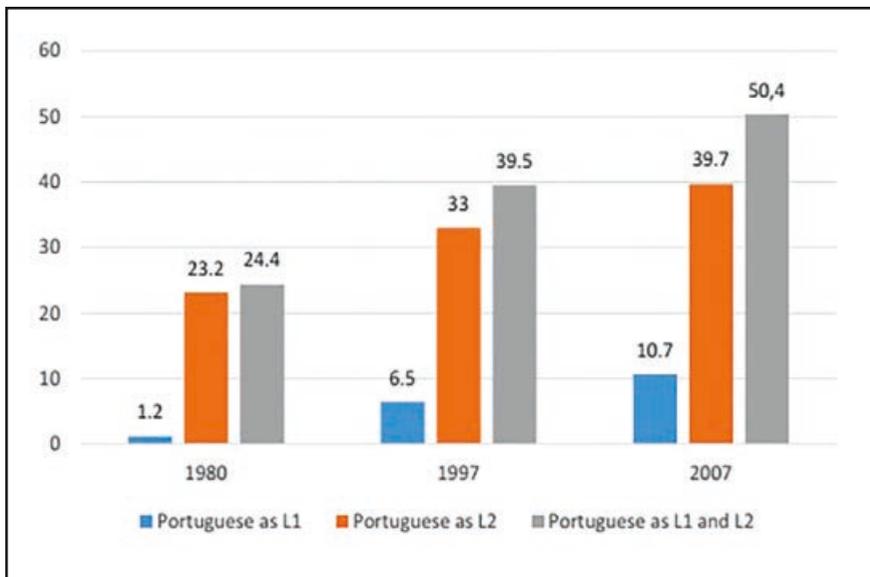
No seu estudo, apresenta dois gráficos baseados nos censos de 1980, 1997 e 2007 que sintetizam a expansão do Português – e o conseqüente decréscimo de uso das línguas Banto (Figura 1) – e o estatuto do Português L1 ou L2 (Figura 2) no decurso de 27 anos.

**Figura 1:** Distribuição da população com uma língua Banto como primeira língua em 1980, 1997 e 2007, em Moçambique



Fontes: Chimbutane (2018, 2012); Firmino (2000).

**Figura 2:** Evolução da proporção de falantes de Português como primeira (L1) ou segunda língua (L2) em Moçambique



Fontes: Chimbutane (2018, 2012); Firmino (2000).

Nesse período, como se verifica pela Figura 1, houve um decréscimo de 13,6% no uso das línguas Banto, enquanto o Português como L1 e L2 cresceu, respectivamente, 10,7% e 39,7%, observando-se, ainda, que o número total de falantes do Português já era de 50,4%, o que está, como observa Chimbutane, vinculado a políticas educacionais postas em prática sobretudo na era pós-colonial.

Cabe ressaltar, como lembra Gonçalves (2000, p. 25-26), que nenhuma das línguas Banto tem caráter majoritário ou se distribui por todas as áreas do país: o Macua, a de maior número de utentes, era falada, de acordo com o censo de 1997, por cerca de 26,3% da população. Em Maputo, por exemplo, as mais frequentes são, por ordem decrescente, o Changana, o Tshwa e o Rhonga, segundo dados do INE, 2010 (*apud* PAULA; DUARTE, 2016).

No que concerne mais propriamente ao PM, Chimbutane (2018, p. 102) ressalta que a expressão Português de Moçambique não constitui, como se poderia esperar, uma variedade falada em todo o país ou que haja alguma forma de padrão estabelecida. Em suas palavras:

Embora estudos sistemáticos de correlatos sociolinguísticos de variação ainda não tenham sido realizados, há uma variação dialetal percebida do português falado em Moçambique. Tomando como referência a variedade padrão europeia, o português moçambicano tem sido descrito como um *continuum*

dialetal constituído por uma gradação desde o acroleto passando pelo mesoleto ao basileto (Gonçalves, 2010, 2012). Descrevendo o *status* dos dois extremos deste *continuum*, Gonçalves (2010) afirma que o basileto engloba as variedades faladas por indivíduos com baixos níveis de escolaridade e o acroleto compreende a variedade falada principalmente por indivíduos urbanos com maiores níveis de escolaridade. Embora Dias (2002) use termos diferentes para caracterizar essa variação, ela também se baseia na mesma abordagem de *continuum* pós-crioulo. Dias (2002) sugere que o *continuum* dialetal compreende variedades pidginizadas, mistas e padronizadas, que correspondem aos termos basileto, mesoleto e acroleto de Gonçalves (2010, 2012), embora com algumas diferenças em termos de caracterização.

Como se verifica, é grande a complexidade do PM, cujas características variam também regionalmente, em decorrência das línguas faladas nas diferentes províncias. Deve-se, portanto, considerar os resultados da pesquisa sobre os róticos que ora se apresenta como um meio de tecer hipóteses para estudos futuros com *corpora* que abranjam não só variedades urbanas.

## ASPECTOS TEÓRICO-METODOLÓGICOS

A pesquisa norteia-se pelos pressupostos teórico-metodológicos da Teoria da Variação e Mudança (WEINREICH; LABOV; HERZOG, 1968), com apoio no Programa GoldVarb-X e tem por base amostras selecionadas do *Corpus* Moçambique-Port, organizado por Vieira; Pissurno, e que constam do site CORPORA PORT (Corpora de variedades do Português em análise), organizado por Brandão; Vieira. As entrevistas foram realizadas com indivíduos que vivem na cidade de Maputo, portanto de perfil urbano. Os 18 informantes, na amostra geral, distribuem-se por sexo, três faixas etárias (18-35, 36-55, 56-75 anos) e três níveis de escolaridade (fundamental, médio e superior), levando-se, ainda, em conta, se são utentes do PL1 ou PL2 e o seu grau de conhecimento/uso de outra(s) língua(s) falada(s) em Moçambique. Considere-se que, no país, como já se mencionou, aproximadamente 10,7% dos habitantes falam o Português como L1 e 39,7%, como L2, índices que podem ter aumentado considerando o Censo de 2017, ainda em processamento.

Para a presente análise distribuíram-se os 18 informantes por dois grupos, segundo o estatuto do Português (L1, com 11 informantes, e L2, com 7), e organizaram-se, para cada um deles, duas amostras, uma referente ao contexto inicial de vocábulo (como em *roça*), outra, ao intervocálico (como em *carroça*), controlando-se variáveis sociais e estruturais. Por conta do número discrepante de informantes em cada um desses grupos, e para minimizar o não preenchimento

de todas as células sociais, amalgamaram-se as variáveis *nível de escolaridade e faixa etária*. Para dar ideia da distribuição dos róticos nas amostras de PL1 e PL2<sup>4</sup>, apresentam-se as Tabelas 1 e 2, a seguir, que também contêm o número de ocorrências de tepes (o valor de aplicação) em contraposição ao de vibrantes alveolares.

**Tabela 1:** Índices relativos ao tepe na fala de informantes de português L1

R FORTE: (1) INICIAL DE VOCÁBULO e (2) INTERVOCÁLICO						
Escol./ Sexo/ Idade	1-Fundamental		2-Médio		3-Superior	
	H	M	H	M	H	M
<b>A</b> <b>18-35</b>		<b>A1m5P</b> (1) 35/36 = 97% (2) 34/37 = 97%	<b>A2h5P</b> (1) 20/27 = 74% (2) 14/18 = 78%		<b>A3h6P</b> (1) 23/51 = 45% (2) 2/11 = 18%	<b>A3m6P</b> (1) 5/46 = 11% (2) 0/23 = 0%
<b>B</b> <b>36-55</b>			<b>B2h5P</b> (1) 18/23 = 78% (2) 9/24 = 37,5%	<b>B2m5P</b> (1) 42/83 = 51% (2) 28/64 = 44%	<b>B3h6P</b> (1) 2/21 = 10% (2) 1/10 = 10%	<b>B3m5P</b> (1) 20/52 = 38% (2) 19/37 = 51%
<b>C</b> <b>56-75</b>		<b>C1m4P</b> (1) 43/67 = 64% (2) 25/64 = 39%			<b>C3h4P</b> (1) 16/66 = 24% (2) 5/28 = 18%	<b>C3m6P</b> (1) 40/84 = 48% (2) 13/74 = 18%

**Tabela 2:** Índices relativos ao tepe na fala de informantes de português L2

R FORTE (1) INICIAL DE VOCÁBULO e (2) INTERVOCÁLICO						
Escol./ Sexo/ Idade	1-Fundamental		2-Médio		3-Superior	
	H	M	H	M	H	M
<b>A</b> <b>18-35</b>	<b>A1h5O</b> (1) 17/20 = 85% (2) 6/8 = 75%			<b>A2m5O</b> (1) 24/27 = 89% (2) 33/39 = 85%		
<b>B</b> <b>36-55</b>	<b>B1h5O</b> (1) 8/12 = 67% (2) 4/8 = 50%	<b>B1m5O</b> (1) 6/17 = 35% (2) 5/5 = 100%				
<b>C</b> <b>56-75</b>	<b>C1h7O</b> (1) 53/61 = 87% (2) 40/50 = 80%		<b>C2h5O</b> (1) 21/33 = 64% (2) 7/11 = 64%	<b>C2m5O</b> (1) 23/26 = 88% (2) 9/10 = 90%		

<sup>4</sup> Os códigos que identificam os informantes indicam faixa etária (A, B e C, respectivamente, 18-35, 36-55 e 56-75 anos), nível de escolaridade (1, 2 e 3, respectivamente, fundamental, médio e superior), sexo (H: homem; M: mulher); estatuto do Português (P: L1; O: L2).

## O R NAS DIFERENTES AMOSTRAS

As tabelas 1 e 2 apresentadas na seção anterior demonstram o predomínio do tepe e da vibrante alveolar, variantes [+ant], nas amostras levantadas, embora haja significativa diferença na performance dos falantes de PL1, em que a vibrante alveolar é mais frequente e do PL2, em que predomina o tepe. Porém, além delas, registraram-se, variantes [-ant] (as fricativas velar, uvular e glotal), com baixos índices: 4,3% e 1,8% no PL1, respectivamente, em contexto inicial e intervocálico; 3% e 0,8%, na mesma ordem contextual, no PL2. Desconsideraram-se esses dados na análise, obtendo-se os índices expostos nas Tabelas 3 e 4, a seguir.

Todos os indivíduos produzem o tepe, que demonstra, assim, ser uma variante em franca competição com a vibrante alveolar nos dois contextos. Apenas uma informante (A3m6P, cf. Quadro 1) não o produziu na posição intervocálica, embora o tenha realizado no contexto inicial de vocábulo (11%). Não se conhece registro do tepe “relativo ao R forte pré-vocálico” no PE; no PB, em que não é uma opção usual, é registrado, em algumas áreas de colonização não portuguesa, como, por exemplo, na Região Sul – em Panambi, de colonização alemã, e em Flores da Cunha, de colonização italiana (BRESCANCINI; MONARETTO, 2008).

Observe-se, ainda, que, no PL1, em que predomina a vibrante alveolar, em contexto inicial (cf. Tabela 3), a diferença entre essa variante e o tepe é de apenas cinco pontos percentuais, embora no contexto intervocálico a diferença aumente para 23%. No PL2, ao contrário, a diferença é de mais de 50% em ambos os contextos. Assim, a vibrante é preterida em ambas as posições na fala dos informantes de PL2, suplantada por altos índices de tepe – 77,6% e 79,4% –, enquanto é recorrente na dos de PL1, com índices majoritários de 52,5%, em início de vocábulo, e de 61,5%, entre vogais. Apesar da presença também expressiva do tepe entre esses falantes, tal diferença para aqueles que têm o Português como PL2 poderia ser indício de que falantes de PL1, em determinadas situações, se dão conta da possibilidade de oposição fonológica no contexto intervocálico, o que não parece ocorrer com os falantes de PL2.

**Tabela 3:** Português de Moçambique como L1 – Variantes de R para análise

Variantes de R forte entre falantes de PL1				
	Inicial de vocábulo (roça)		Intervocálico (carro)	
	Nº	%	Nº	%
Tepe	264/556	47,5%	150/390	38,5
<b>Vibrante alveolar</b>	292/556	52,5	240/390	<b>61,5</b>

**Tabela 4:** Português de Moçambique como L2 – Variantes de R para análise

Variantes de R forte entre falantes de PL2				
	Inicial de vocábulo (roça)		Intervocálico (carro)	
	Nº	%	Nº	%
<b>Tepe</b>	152/196	77,6	104/131	<b>79,4</b>
Vibrante alveolar	44/196	22,4	27/131	20,6

Na seção seguinte, apresentam-se, de forma contrastiva, os resultados das análises multivariadas relativas a L1 e a L2, por contexto.

## R EM CONTEXTO INICIAL DE VOCÁBULO

No Quadro 4, sintetizam-se os principais resultados das análises. Nele, fica clara a diferença no uso do tepe entre PL1 e PL2, assim como indicaram os índices gerais. Entre os falantes PL1, o tepe alterna com a vibrante de forma equilibrada, beirando a faixa dos 50%, e a probabilidade de aplicação da regra é neutra, com *input* de .49. Já os falantes de PL2 implementam preferencialmente o tepe, com *input* .83 e 77,6% de uso.

**Quadro 4:** Síntese dos resultados das análises de R forte inicial de vocábulo (roça) no Português de Moçambique L1 e L2

PM L1		PM L2	
<b>R INICIAL DE VOCÁBULO</b>		<b>R INICIAL DE VOCÁBULO</b>	
<b>Valor de aplicação:</b> <b>Tepe (x vibrante alveolar)</b> <b>Tepe: 47,5%</b>		<b>Valor de aplicação:</b> <b>Tepe (x vibrante alveolar)</b> <b>Tepe: 77,6%</b>	
<b>VARIÁVEIS SELECIONADAS</b> <i>Tonicidade da sílaba</i> <i>Nível de escolaridade/faixa etária</i> <i>Número de sílabas do vocábulo</i>		<b>VARIÁVEIS SELECIONADAS</b> <i>Tonicidade da sílaba</i> <i>Nível de escolaridade/faixa etária</i> <i>Classe do vocábulo</i>	
Input: .49	Significância: .039	Input: .83	Significância: .022

Entre as variáveis estruturais selecionadas na análise multivariada, estão aquelas de natureza linguística: *tonicidade da sílaba*, tanto para o PL1 quanto para o PL2; *número de sílabas do vocábulo*, para PL1; e *classe do vocábulo* para PL2, as duas últimas não selecionadas na análise de Brandão; De Paula (2018) que levou em conta os dois grupos em conjunto (cf. Quadros 2 e 3 da seção 1).

Destaca-se, ainda, a variável social composta *nível de escolaridade/faixa etária*, devendo-se atentar que, nessa relação, cada fator equivale a um ou dois falantes (cf. Tabelas 1 e 2) e algumas células não estão preenchidas. Tal expediente metodológico permitiu não só verificar, até certo ponto, a performance de cada falante sem deixar de contemplar a atuação dessas variáveis, mas também salientar a complexidade sociolinguística de Moçambique.

No que tange aos falantes de PL1, um olhar atento sobre tal variável, explanada na Tabela 5, demonstra que os falantes com nível superior de escolaridade são os que menos implementam o tepe, com pouca diferença entre as faixas etárias 1, 2 e 3 (P. R. .27, .29 e .35). Enquanto isso, os falantes com menos escolaridade (níveis fundamental e médio) preferem o tepe (P. R. .97 e .64, no ensino fundamental; e .80 e .58, no ensino médio). Os da faixa etária 1, no extremo de um contínuo de acesso à educação escolar, chegam a um percentual de uso de tepe de 97,2% e P. R. .97. Por tudo isso, o aumento da escolaridade parece claramente inibir o uso do tepe no lugar da vibrante.

**Tabela 5:** Atuação da variável compósita *nível de escolaridade/faixa etária* para a implementação do tepe em contexto inicial de vocábulo (roça) no PM L1

<b>Variável compósita <i>nível de escolaridade/faixa etária</i> no PM L1</b>			
<b>Fatores</b>	<b>Apl./Nº</b>	<b>%</b>	<b>P.R.</b>
<b>Nível Fund-18-35 anos</b>	35/36	97,2	<b>.97</b>
<b>Nível Fund-56-75 anos</b>	43/67	64,2	<b>.64</b>
<b>Nível Méd-18-35 anos</b>	20/27	74,1	<b>.80</b>
<b>Nível Méd-36-55 anos</b>	60/106	56,6	<b>.58</b>
<b>Nível Sup-18-35 anos</b>	28/97	28,9	.27
<b>Nível Sup-36-55anos</b>	22/73	30,1	.29
<b>Nível Sup-56-75 anos</b>	56/150	37,3	.35
Input: .49		Significância: .039	

Foram também selecionadas, nessa amostra, as variáveis linguísticas *tonicidade da sílaba* e *número de sílabas do vocábulo*, tendo sido o fator relevante para o uso do tepe, a sílaba pretônica (P. R. .55), enquanto a tônica mostrou inibi-lo (P. R. .37), como se observa na Tabela 6.

**Tabela 6:** Atuação da variável *tonicidade da sílaba* para a implementação do tepe em contexto inicial de vocábulo (roça) no PM L1

<b>Variável <i>tonicidade da sílaba</i> no PM L1</b>			
<b>Fatores</b>	<b>Apl./Nº</b>	<b>%</b>	<b>P.R.</b>
<b>Pretônica</b>	200/389	51,4	<b>.55</b>
<b>Tônica</b>	64/167	38,3	.37
Input: .49		Significância: .039	

Quanto à extensão do vocábulo (Tabela 7), são as palavras com apenas uma sílaba (P. R. .69) que parecem favorecer o tepe. De qualquer forma, neste caso, não fica claro que palavras mais extensas diminuam o uso do tepe, pois não só a diferença entre os pesos relativos dos fatores é pequena como também os termos com quatro sílabas ou mais apresentam peso relativo acima da neutralidade (.58):

**Tabela 7:** Atuação da variável *número de sílabas do vocábulo* para a implementação do tepe em contexto inicial de vocábulo (roça) no PM L1

<b>Variável número de sílabas no PM L1</b>			
<b>Fatores</b>	<b>Apl./Nº</b>	<b>%</b>	<b>P.R.</b>
<b>Uma sílaba</b>	11/18	61,1	.69
<b>Duas sílabas</b>	63/174	36,2	.44
<b>Três sílabas</b>	104/201	47,7	.47
<b>Quatro ou mais sílabas</b>	86/146	58,9	.58
Input: .49		Significância: .039	

Na análise dos falantes de PL2, como visto, duas variáveis foram selecionadas em comum com PL1. Quanto à *escolaridade/faixa etária* (Tabela 8), entretanto, não é possível verificar a sua atuação da mesma forma que no PL1, visto que nenhum dos falantes de L2 da amostra geral tem nível superior.

**Tabela 8:** Atuação da variável compósita *nível de escolaridade/faixa etária* para a implementação do tepe em contexto inicial de vocábulo (roça) no PM L2

<b>Variável compósita nível de escolaridade/faixa etária no PM L2</b>			
<b>Fatores</b>	<b>Apl./Nº</b>	<b>%</b>	<b>P.R.</b>
<b>Nível Fund-18-35 anos</b>	17/20	85	.55
<b>Nível Fund-36-55 anos</b>	14/29	48,3	.25
<b>Nível Fund-56-75 anos</b>	53/61	86,9	.60
<b>Nível Méd-18-35 anos</b>	24/27	88,9	.75
<b>Nível Méd-56-75 anos</b>	44/59	74,6	.37
Input: .83		Significância: .022	

Já quanto à tonicidade da sílaba, também as pretônicas favorecem o uso do tepe (P. R. .60), como entre os falantes PL1 (P. R. .55). Dessa forma, nos dois grupos, parece que a fragilidade do contexto átono está relacionada com o tepe, enquanto a sílaba tônica seria mais suscetível à vibrante múltipla.

**Tabela 9:** Atuação da variável *tonicidade da sílaba* para a implementação do tepe em contexto inicial de vocábulo (roça) no PM L2

<b>Variável <i>tonicidade da sílaba</i> no PM L2</b>			
<b>Fatores</b>	<b>Apl./Nº</b>	<b>%</b>	<b>P.R.</b>
<b>Pretônica</b>	129/148	87,2	<b>.60</b>
<b>Tônica</b>	23/48	47,9	.20
Input: .83		Significância: .022	

Além das duas variáveis em comum com PL1, também foi selecionada a *classe do vocábulo*. Esse quesito, que é relevante apenas para os falantes PL2, aponta os *verbos* como a classe em que o tepe é mais usado, em face das demais (P. R. .40), como se demonstra a seguir.

**Tabela 10:** Atuação da variável *classe do vocábulo* para a implementação do tepe em contexto inicial de vocábulo (roça) no PM L2

<b>Variável <i>classe do vocábulo</i> no PM L2</b>			
<b>Fatores</b>	<b>Apl./Nº</b>	<b>%</b>	<b>P.R.</b>
<b>Verbo</b>	64/71	90,1	.66
<b>Nome e outras</b>	88/125	70,4	.40
Input: .83		Significância: .022	

## R EM CONTEXTO INTERVOCÁLICO

No contexto intervocálico, é notória a presença do tepe como opção para a realização do R “forte”, entre falantes tanto de PL1 quanto de L2, o que demonstra ser essa uma característica muito proeminente do Português falado em Moçambique, que o distingue do PE e do PB.

Com os resultados abaixo, ratifica-se, e fica ainda mais evidente, a diferente performance dos falantes de PL1 e de PL2. Os primeiros aplicam o tepe em 38,5% dos casos e os últimos, em 79,4%. O primeiro desses índices, embora equivalente à metade do segundo, é expressivo, visto se tratar de um contexto de oposição fonológica. Assim, no PM em geral, tal oposição parece ter algum nível de instabilidade, o que voltará a ser comentado adiante.

Entre os falantes de L1, mostrou-se relevante apenas a variável composta *nível de escolaridade/faixa etária* e, entre os de L2, uma variável social, *sexo*, além do *contexto subsequente* ao R, resultados expostos no Quadro 5.

**Quadro 5:** Principais resultados das análises de R intervocálico (carro) no Português de Moçambique falado como L1 e como L2

PM L1	PM L2
<p><b>R INTERVOCÁLICO</b></p> <p>Valor de aplicação: Tepe (x vibrante alveolar) Tepe: 38,5%</p> <p><b>VARIÁVEL SELECIONADA</b> <i>Nível de escolaridade/faixa etária</i></p> <p>Input: .36                      Significância: .000</p>	<p><b>R INTERVOCÁLICO</b></p> <p>Valor de aplicação: Tepe (x vibrante alveolar) Tepe: 79,4%</p> <p><b>VARIÁVEIS SELECIONADAS</b> <i>Sexo</i> <i>Contexto subsequente</i></p> <p>Input: .86                      Significância: .017</p>

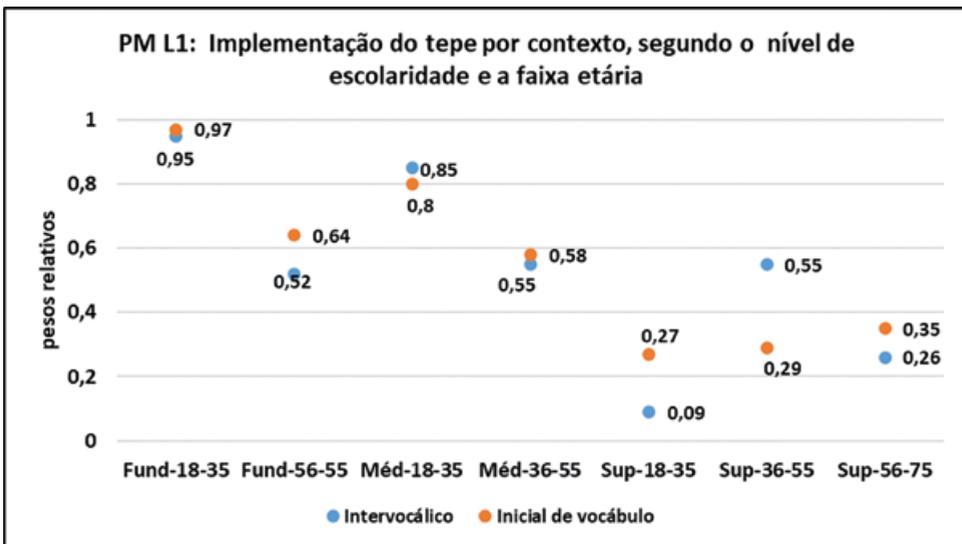
Apesar de a preferência pela vibrante por parte dos falantes de PL1 (nos índices gerais) aproximá-los da norma mais conservadora do PE, a atuação da escolaridade para que o tepe seja evitado pode indicar que a consciência fonológica de oposição entre os róticos no contexto intervocálico dependa, no PM, de maior nível de escolaridade. Como se verifica na Tabela 9, assim como no contexto inicial, os falantes jovens que cursaram apenas o nível fundamental usam o tepe acima dos 90%, nesse caso com P. R. de .95, enquanto os menores índices, P. R. .09 e .26, encontram-se, na maioria, entre os falantes que concluíram o nível superior. Também os jovens com nível médio parecem preferir o tepe de forma mais evidente que os mais velhos com alta escolaridade, com índice de P. R. .85.

**Tabela 11:** Atuação da variável compósita *nível de escolaridade/faixa etária* para a implementação do tepe em contexto intervocálico (carro) no PM L1

<b>Variável compósita <i>nível de escolaridade/faixa etária</i> no PM L1</b>			
<b>Fatores</b>	<b>Apl./Nº</b>	<b>%</b>	<b>P.R.</b>
Nível Fund-18-35 anos	34/37	91.9	.95
Nível Fund-56-75 anos	25/64	39,1	.52
Nível Méd-18-35 anos	14/18	77.8	.85
Nível Méd-36-55 anos	37/88	42	.55
Nível Sup-18-35 anos	2/34	5.9	.09
Nível Sup-36-55anos	20/47	42.6	.55
Nível Sup-56-75 anos	18/102	17,6	.26
Input: .436		Significância: .000	

É interessante comparar os resultados do PL1 para o contexto intervocálico aos obtidos na posição inicial de vocábulo, expostos na Figura 3. Os índices relativos aos dois contextos são semelhantes, só havendo discrepância na fala dos indivíduos da faixa etária média de nível superior (em realce com parênteses no gráfico), os quais destoam da tendência geral das amostras, com índices de tepe em contexto intervocálico que superam os do contexto inicial de sílaba, o que se deve à diferente performance dos informantes do sexo masculino e feminino.

**Figura 3.** Implementação do tepe segundo o contexto e o nível de escolaridade/faixa etária no PM L1



No PL2 (cf. Tabelas 12 e 13), a variável *escolaridade/faixa etária* não foi selecionada para o contexto intervocálico, no entanto, mais uma vez, foi uma variável de cunho social – o *sexo* – a que se mostrou mais representativa. As mulheres implementaram mais o tepe – aplicando essa variante com P. R. .67 –, enquanto os homens apresentaram P. R. .37.

**Tabela 12:** Atuação da variável *sexo* para a implementação do tepe em contexto intervocálico (carro) no PM L2

Variável <i>sexo</i> no PM L2			
Fatores	Apl./Nº	%	P.R.
Homem	57/77	74	.37
Mulher	47/54	87	.67
Input: .86		Significância: .017	

Em segundo lugar, aparece selecionada a variável *contexto subsequente*, com resultado que indica as vogais [ɛ e i] (P. R. .82) como as principais condicionadoras do tepe, que com elas partilha o traço coronal. A vogal central [a] e as vogais posteriores [ɔ o u] têm índices muito abaixo delas, com peso relativo de .30 e .22, respectivamente:

**Tabela 13:** Atuação da variável *contexto subsequente* para a implementação do tepe em contexto intervocálico (carro) no PM L2

Variável <i>contexto subsequente</i> no PM L2			
Fatores	Apl./Nº	%	P.R.
[a]	27/39	69,2	.30
[ɛ e i]	50/52	96,2	<b>.82</b>
[ɔ o u]	27/40	67,5	.22
Input: .86		Significância: .017	

Conforme se pode verificar pela explanação dos resultados relativos a ambos os contextos, a despeito da relevância de fatores de natureza estrutural, fica clara a significativa atuação de variáveis de cunho social, como a da variável compósita *faixa etária/nível de escolaridade*, o que corrobora a observação de Chimbutane de que o PM, em relação ao PE, constitui um *continuum* dialetal.

## DISCUSSÃO

Os resultados expostos nas seções anteriores demonstram, de um lado, a complexidade de que se reveste a emergente variedade moçambicana do Português, tanto como L1 quanto como L2, de outro, a importância da variável *nível de escolaridade/faixa etária*, entre outras que concorrem para a variação observada. Além disso, os resultados parecem ratificar e exemplificar as observações de Chimbutane (2018), quando este faz uma apreciação crítica das políticas educacionais nas eras colonial e pós-colonial responsáveis pelo “atual estado de nativização do Português em Moçambique” (p. 95). Vale a pena sintetizar suas observações.

Segundo o autor, no primeiro período da era colonial (séc. XV a 1926), a educação de nativos era feita por missionários. Só em 1930 teve início a educação colonial propriamente dita, numa associação do Estado português com a Igreja Católica, alavancada pelo relativo crescimento do número de colonos portugueses e pela necessidade de controlar a educação dos nativos, o que

levou a um sistema discriminatório: o “ensino oficial”, voltado para os filhos dos colonos e dos assimilados e que visava à formação de uma elite a serviço do Estado, e o “ensino rudimentar” direcionado aos nativos, feito por missionários e destinado a “equipar os nativos com conhecimentos rudimentares e valores morais” (p. 95-96). Nessa fase, o uso das línguas africanas na educação era proibido, o que passou a mudar a partir de 1960, entre outros aspectos, para facilitar nas escolas primárias a aprendizagem do português. Em 1974, um ano antes da independência, o analfabetismo era da ordem de 93% entre os indivíduos de mais de 7 anos. Tudo isso, reduziu “o nível de contato entre o Português e as línguas locais, restringindo a nativização daquela língua no Moçambique colonial” (p. 97-98).

Com a independência, em 1975, o Português foi declarado a única língua oficial e, por questões político-ideológicas, a língua da unidade nacional, mantendo-se as línguas africanas no mesmo patamar da era colonial, o que só veio a mudar muito recentemente. A estratégia de difundir o Português acabou por determinar a construção de escolas e a alfabetização de adultos, o que redundou na redução para 72% em 1980 e cerca 49% em 2007 da taxa de analfabetismo (p. 98).

Chimbutane (p. 98-100), com base em Diniz (1995), menciona duas fases no que se relaciona a medidas adotadas no ensino de Português entre 1975 e 1990: na primeira, que vigoraria até 1979, não haveria clareza sobre o ensino de Português como L2 nem sobre o papel das L1s dos alunos; na segunda (1979-1990), houve a reafirmação do Português como língua da unidade nacional, o incentivo a seu uso massivo, mas também, a par do objetivo de desenvolver uma variedade nativizada, criar condições para sua subsequente standardização. Chimbutane acrescenta uma terceira fase (1990 aos dias atuais) iniciada com a inclusão na Constituição de artigo que estabelece que o Estado promova o desenvolvimento e aumento do uso de línguas africanas na vida pública, inclusive na educação. Desse modo, se teriam ampliado as condições para maior contato entre o Português e as demais línguas, e, conseqüentemente, maiores chances de sua nativização.

As considerações de Chimbutane e a noção de *continuum* dialetal são importantes para a compreensão do que ocorre no PM com os róticos e, ainda, com outros fenômenos variáveis em que o Estatuto do Português e o contato com outras línguas se mostraram relevantes.

Gonçalves (2010) afirma que entre os falantes que têm como L1 o Changanha (uma das cerca de 26 línguas faladas em Moçambique, predominante em

Maputo, o local da recolha de dados), o uso do tepe no PL2 seria um processo de hipercorreção, uma vez que o Changana apresenta apenas um rótico, a vibrante alveolar, que aparece frequentemente em palavras como *areia* e *herói* – nas quais o R português é “fraco” – ao mesmo tempo em que o tepe é usado em termos como *carro* e *morrer* (2010: 42), em que o R esperado seria o “forte”.

A hipótese da hipercorreção parece não se sustentar pelo alto índice de uso do tepe no PL2 e mesmo no PL1, em menor escala. Embora ainda não se tenha analisado o R “fraco” intervocálico, a audição das fitas demonstrou que também a vibrante alveolar é empregada nesse contexto, embora não de forma tão frequente. Vale a pena mencionar o depoimento da informante A2m5O (cf. Quadro 2) em relação ao contexto de oposição fonológica:

*DOC 1: “e uma palavra assim como “caro”, “carro”, quando for “um automóvel está caro”, como é que você diz? o..*

*L: o chapa tá car[r]o*

*DOC 2: “e se você quiser falar “carro,” como é que fala?”*

*L: O chapa tá ca[r]o*

*DOC 1: “e uma palavra assim como “caro”, “carro”, quando for “um automóvel está caro”, como é que você diz? o...”*

*L: aquele ca[r]o está ca[r]o ((muitos risos))*

*Doc 2: isso é mais comum aqui, né?*

*L: sim, é assim mesmo...porque o car[r]o é pro automóvel e o ca[r]o já é para o preço que é elevado*

*Doc2: entendi, entendi*

A observação da performance de falantes de L1 também contradiz a hipótese. Tomando como exemplo o emblemático *carro*, verifica-se que, na fala de A1m5P, as oito ocorrências desse vocábulo foram produzidas com tepe, enquanto na de A3h6P, entre 5 ocorrências, 4 foram de vibrante alveolar e uma de tepe.

Ngunga; Simbini (2012, p. 29), em sua gramática do Changana, afirmam haver apenas a vibrante alveolar no seu sistema fonológico, apresentando como exemplos *kurila* (chorar); *rito* (voz); *murimi* (camponês), exatamente os contextos aqui focalizados. No entanto, não há qualquer observação sobre possíveis casos de variação.

Assim, pode-se formular a hipótese de que a variação encontrada no PM como um todo e, sobretudo no PM L2, deve-se não só ao contato com as línguas

faladas em Moçambique, como o Changana, que predomina em Maputo, a área da pesquisa, mas também com o PE.

De um lado, uma língua (como o Changana, por exemplo) que, até prova em contrário, apresenta categoricamente a vibrante alveolar; de outro, uma variedade do Português em que:

- (a) os pares mínimos, em contexto de oposição fonológica, têm baixíssima produtividade;
- (b) o tepe tem uma maior distribuição do que as demais variantes: ocorre em contexto intervocálico (correspondente a R [+ant]), e predomina, ainda, em ataque complexo e em coda silábica interna e externa.

Além disso, a diferença articulatória entre um tepe e uma vibrante alveolar consiste basicamente, como afirma Lindau (1985), entre uma batida e mais de uma batida nos alvéolos, o que pode não ser saliente para o indivíduo que adotou o Português (BRANDÃO; DE PAULA, 2018).

Assim, tudo indica que, no plano fonológico do PM, haja apenas um R e que a variação dependa de fatores sobretudo de natureza social, embora uma visão global do comportamento dos róticos nessa variedade ainda dependa dos resultados de pesquisas recém-iniciadas e que têm por foco o R em *onset* complexo (como em *preso, quatro*) e o R intervocálico (em contextos como *caro, senhora e espera*).

## CONSIDERAÇÕES FINAIS

Embora as análises apresentadas não tenham seguido os princípios de ortogonalidade requeridos em análises sociolinguísticas clássicas, a opção metodológica adotada parece ter contribuído para enfatizar a complexidade que caracteriza a sociedade moçambicana, em que diversas línguas nativas concorrem para a definição da identidade sociolinguística dos indivíduos. Nesse sentido, futuras investigações têm, entre outros aspectos, de basear-se em amostras que levem em conta indivíduos com diferentes perfis, em outras áreas urbanas e em áreas rurais, com baixo ou nulo nível de escolaridade, usuários de diferentes línguas Banto, como, aliás, ressalta Chimbutane (2018). Só assim se poderá comprovar a hipótese aqui defendida de que na gramática do PM, a despeito da variação, há apenas um fonema rótico.

## REFERÊNCIAS

BOUCHARD, Marie-Eve. *Linguistic variation and change in the Portuguese of São Tomé*. Doctoral dissertation. New York University, 2017.

BRANDÃO, Sílvia Figueiredo. Apagamento de R em coda externa em duas variedades africanas do Português. *Diadorim*, v.20, n. especial, p. 390 - 408, 2018.

BRANDÃO, Sílvia Figueiredo; DE PAULA, Alessandra. Róticos nas variedades urbanas santomense e moçambicana do Português. In: BRANDÃO, Sílvia Figueiredo (Org). *Duas variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher, 2018, p. 95-118.

BRANDÃO, Sílvia Figueiredo; DE PAULA, Alessandra. Um estudo comparativo preliminar sobre os róticos em duas variedades africanas do Português. Comunicação apresentada ao Encontro Intermediário do GT de Sociolinguística da ANPOLL, Porto Alegre, 22-24 de novembro de 2017.

BRANDÃO, Sílvia Figueiredo; PESSANHA, Davi Bretas dos Santos; PONTES, Stefany de Paulo; CORRÊA, Monique Oliveira. Róticos na variedade urbana do Português de São Tomé. *Papia*, v. 27, n. 2, p. 191-213, 2017.

BRESCANCINI, Cláudia; MONARETTO, Valéria Neto de Oliveira. Os róticos no sul do Brasil: panorama e generalizações. *SIGNUM: Estud. Ling.*, v.11, n. 2, p. 51-66, 2008.

CHIMBUTANE, Feliciano. Portuguese and African languages in Mozambique. In: LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de (Orgs.). *The portuguese language continuum in Africa and Brazil*. Amsterdam: John Benjamins, 2018, p. 89-110.

CHIMBUTANE, Feliciano. *Panorama linguístico de Moçambique: Análise dos dados do III recenseamento geral da população e habitação de 2007*. Maputo: Instituto Nacional de Estatística, 2012.

DIAS, Hildizina Norberto. *As desigualdades sociolinguísticas e o fracasso escolar: em direcção a uma prática linguístico-escolar libertadora*. Maputo: Promédia, 2002.

FIRMINO, Gregório. *Situação linguística de Moçambique: Dados do II recenseamento geral da população e habitação de 1997*. Maputo: Instituto Nacional de Estatística, 2000.

GONÇALVES, Perpétua. O Português em África. In: RAPOSO, Eduardo Buzaglo Paiva; NASCIMENTO, Maria Fernanda Bacelar do; MOTA, Maria Antónia Coelho da; SEGURA, Luísa; MENDES, Amália. (Orgs.) *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, 2013. Vol. 1, p. 157-178.

GONÇALVES, Perpétua. *A génese do Português de Moçambique*. Lisboa: Imprensa Nacional/Casa da Moeda, 2010.

HAGEMEIJER, Tjerk. Prefácio. In: BRANDÃO, Silvia Figueiredo (Org.) *Duas variedades africanas do Português: variáveis fonético-fonológicas e morfosintáticas*. São Paulo: Blucher, 2018. p. 8-9.

LADEFOGED, Peter; MADDIESON, Ian. *The sounds of world's languages*. Oxford: Blackwell, 1996.

LINDAU, Mona. The story of /r/. In: FROMKIM, Victoria A. (Org.) *Phonetic Linguistics: essays in honor of Peter Ladefoged*. Orlando: Academic Press, 1985. p. 157-168.

MATEUS, Maria Helena Mira; D'ANDRADE, Ernesto. *The phonology of Portuguese*. Oxford: University Press, 2000.

NGUNGA, Armindo; SIMBINI, Madalena Cítia. *Gramática descritiva da língua changana*. 2012. Maputo: Centro de Estudos Africanos (CEA)/UEM, 2012.

PAULA Ronaldo Rodrigues de; DUARTE, Fábio Bonfim. Diversidade linguística em Moçambique. In: LEITE, Ilka Boaventura; SEVERO, Cristine Gorski (Orgs.) *Kadila: culturas e ambientes; diálogos Brasil-Angola*. São Paulo: Blucher, 2016, p. 343-362.

PEREIRA, Rodrigo; HAGEMEIJER, Tjerk; FREITAS Maria João. Consoantes róticas e variação no português de São Tomé. *Revista da Associação Portuguesa de Linguística*, n.4, p. 206-224.

VELOSO, João. The english R coming! The never ending story of Portuguese rhotics. In: SIMÕES, A.; BARREIRO A.; SANTOS, D.; SOUSA-SILVA, R.; TAGNIN, E. O. (Eds.) *Linguística, Informática e Tradução: Mundos que se cruzam*. Oslo, Studies in Language, v. 7, n.1, p. 323–336. 2015. Disponível em: <<http://www.journals.uio.no/osla>>. Acesso em: 10 fev. 2016.

VIEIRA, Silvia Rodrigues; BRANDÃO, Silvia Figueiredo. CORPORAPORT: Variedades do Português em análise. Rio de Janeiro: Faculdade de Letras-UFRJ. Disponível em: <[www.corporaport.letas.ufrj.br](http://www.corporaport.letas.ufrj.br)> Acesso em: 10 jan. 2019.

VIEIRA, Silvia Rodrigues; PISSURNO, Karen Cristina da Silva (Orgs.). *Corpus Moçambique-PORT*. Rio de Janeiro: Faculdade de Letras/UFRJ. Disponível em: <[www.corporaport.letas.ufrj.br](http://www.corporaport.letas.ufrj.br)> Acesso em: 10 jan. 2019.

WEINREICH, Uriel, LABOV, William; HERZOG, Marvin. Empirical foundations for a theory of linguistic change. In: LEHMANN, Winfred P.; MALKIEL, Yakov (orgs.) *Directions for historical linguistics*. Austin: University of Texas Press, 1968, p. 97-195.



# STILL ON RTHOTICS IN PORTUGUESE IN MOZAMBIQUE

*Silvia Figueiredo Brandão (UFRJ/CNPq)*  
*Alessandra de Paula (UERJ/FAPERJ)*

In this chapter, we focus on rhotics in the urban variety of Portuguese of Mozambique (PM) in the light of theoretical and methodological assumptions of the Theory of Variation and Change (WEINREICH, LABOV; HERZOG, 1968) in order to discuss structural and social factors, which condition the use of variants of R in pre-vocalic position – as in *rosa* and *carro* – in the speech of individuals who use Portuguese as L1 (PL1) and L2 (PL2).

PM occurs in an environment of intense multilingualism. As pointed out by Chimbutane (2018), in the last 30 years, there was a decrease of 13.6% in the usage of Bantu languages, while Portuguese as L1 and L2 grew, respectively, 10.7% and 39.7%, in addition to the fact that the total number of Portuguese speakers reached 50.4% in 2007, which is, as Chimbutane observes, linked to educational policies practiced especially in the postcolonial era. Other researches also point out that none of Bantu languages is the majority one or is distributed throughout the country. In this sense, Macua, that has the largest number of users, was reported by 26.3% of the population (Censo1997), but in Maputo, where the present survey was conducted, the most frequent are, in descending order, Changana, Tshwa and Rhonga (INE 2010). As shown, the complexity of PM is great and its characteristics vary also regionally, due to languages spoken

in the capital and in different provinces. Therefore, we should consider the results of this research as a means of formulating hypotheses for future studies using *corpora* that include not only urban varieties.

Brandão; De Paula (2018) examined the rhotics, in pre and post-vocalic contexts, with the same Mozambican informants who contributed to this analysis. However, they focused the individuals on a different perspective, considering PL1 and PL2 speakers together. The results have motivated the continuity of the study, because the relevance of the variable *Portuguese Status* (L1 or L2) was demonstrated, especially in the intervocalic position. In this context, the tap was the most frequent variant for PL2 speakers. The results seemed to show that, in the grammar of most PM speakers, the [+ ant] and [-ant] rhotics, which act in the context of phonological opposition, both in European Portuguese (PE) and Brazilian Portuguese (PB), have suffered a process of neutralization in PM.

In this stage of analysis, Portuguese L1 speakers are focused in contrast to L2 ones. So, we focus the called “strong R” in the initial and *intervocalic positions*, seeking, among other objectives, to verify whether the factors that contribute, in each case, for the implementation of the tap are the same, and also whether the two rhotic phonemes are represented in the grammar of the upper-level individuals of these two social segments.

The variationist analyzes were performed with the support of GoldVarb-X Program and is based on samples selected from *Corpus Mozambique-Port*. The interviews were conducted with individuals living in the city of Maputo, thus having an urban profile. The 18 informants, from the general sample, were distributed by sex, three age ranges (18-35, 36-55, 56-75 years) and three levels of education (lower secondary, upper secondary and higher), also considering whether they are PL1 or PL2 users and their level of knowledge and use of other(s) language(s) spoken in Mozambique.

The results show a clear difference in the use of tap between PL1 and PL2. Among PL1 speakers, the tap alternates with the trill in a balanced way, bordering the 50% range, and the probability of applying the rule is neutral, with *input* .49. The PL2 speakers preferentially implement the tap, with *input* .83 and 77.6% of use.

Among the structural variables selected in the multivariate analysis, stand out *syllable tonicity* for both PL1 and PL2; *number of syllables*, for PL1; and *word class*, for L2.

Regarding PL1 speakers, those with a higher level of education are the ones that least implement the tap, with little difference between the age groups 1, 2

and 3 (P. R. .27, .29 and .35). Meanwhile, speakers with less education (lower and upper secondary) prefer tap (P. R. .97 and .64, in lower level, and .80 and .58, in upper level). The age group 1, starting the *continuum* of schooling, reach a percentage of 97.2% of tap and P. R. .97. For all this, the increase in schooling seems clearly to inhibit the use of the tap in place of the trill.

In intervocalic context, the presence of the tap as an option for the “strong R” among PL1 and PL2 speakers is notorious and it confirms to be a very prominent feature of Portuguese spoken in Mozambique, distinct from PE and PB. In this context, the different performances of PL1 and PL2 speakers become even more evident. The former apply tap in 38.5% of the cases and the latter in 79.4%. The first of these indexes, although equal to half of the second one, is expressive, since it is a context of phonological opposition. Thus, in PM in general, that opposition seems to express some level of instability. Regarding the L1 speakers, only the composite variable *level of education/age group* was relevant and, among the L2 ones, the social variable *sex*, in addition to *subsequent context* to R.

In both contexts, besides the relevance of structural factors, the significant performance of social variables, such the composite variable *level of education/age group*, is clear and it corroborates Chimbutane’s observation that PM, in relation to PE, constitutes a lectal *continuum*.

The results show, on the one hand, the complexity of the emergent Mozambican variety of Portuguese, both as L1 and as L2; on the other, the importance of the variable *level of education/age group*, besides others that contribute for the observed variation. In addition, the results seem to ratify and exemplify Chimbutane’s (2018) observations, when he makes a critical appraisal of educational policies in the colonial and postcolonial eras, which are responsible for the “current state of nativization of Portuguese in Mozambique” (page 95).

Chimbutane points out, in these public policies, three stages as regards the measures adopted in the teaching of Portuguese: (i) the first one, which persisted until 1979, did not clarify the purpose of teaching European Portuguese as L2 nor the role of students’ first language; (ii) the second one, from 1979 to 1990, reaffirmed Portuguese as the language of national unity and incentivized its massive use, for the purpose of to develop a nativized variety and its subsequent standardization; (iii) the third phase, which extends from 1990 on aims to repair the failure of the previous one, with the inclusion in the Constitution of an article which establishes the promotion, by the State, of the development and increase of the use of African languages in life education. Thereby, the conditions for a

higher contact between Portuguese and other languages – and, consequently, more chances of their nativization – would have been increased.

It is possible to formulate the hypothesis that the variation found in the PM, especially in the PM L2, is due not only to the contact with the Bantu languages, as, in the case of Maputo, Changana, but also with the EP, a variety of Portuguese in which: (a) the minimal pairs, in the context of phonological opposition, have very low productivity; and (b) the tap has a larger distribution than the other variants. Moreover, the articulatory difference between a tap and an alveolar trill consists basically, as Lindau (1985) states, in one beat or more than one beat in the alveoli, which may not be salient for the individual who adopted the Portuguese language (BRANDÃO; DE PAULA, 2018). Thus, everything indicates that, in the PM phonological level, there is only one R and that this variation depends on factors mainly of a social nature, although an overview of the behavior of the rhotics in this variety still depends on the results of newly initiated researches which focus the R in complex onset (as in *preso, quatro*) and in the intervocalic R position (in contexts as *caro, senhora* and *espera*).

These analyzes contributed to emphasize the complexity that characterizes the Mozambican society, in which numerous native languages contribute to the definition of the sociolinguistic identity of its individuals. In this sense, future research must, including other aspects, to be based on samples that take into account people with different profiles, in other urban areas and in rural areas, with low or null level of education, users of different Bantu languages, as states Chimbutane (2018). Only then it will be possible to prove the hypothesis here defended, that, despite the variation, in the PM grammar there is only one **rhotic** phoneme.

## REFERENCES

BRANDÃO, Silvia Figueiredo; DE PAULA, Alessandra. Róticos nas variedades urbanas santomense e moçambicana do Português. In: BRANDÃO, Silvia Figueiredo (Org). *Dois variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher, 2018. p. 95-118.

BRANDÃO, Silvia Figueiredo; DE PAULA, Alessandra. Um estudo comparativo preliminar sobre os róticos em duas variedades africanas do Português. Comunicação apresentada ao Encontro Intermediário do GT de Sociolinguística da ANPOLL, Porto Alegre, 22-24 de novembro de 2017.

CHIMBUTANE, Feliciano. Portuguese and African languages in Mozambique. In: LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de (Orgs.). *The portuguese language continuum in Africa and Brazil*. Amsterdam, John Benjamins, 2018. p. 89-110.

LINDAU, Mona. The story of /r/. In: FROMKIM, Victoria A. (org.). *Phonetic Linguistics: essays in honor of Peter Ladefoged*. Orlando: Academic Press, 1985. p. 157-168.



# PARA UMA NORMA-PADRÃO FLEXÍVEL NO CONTEXTO ESCOLAR CONTRIBUIÇÕES DOS ESTUDOS SOCIOLINGUÍSTICOS

*Silvia Rodrigues Vieira (UFRJ)*

### RESUMO

O presente capítulo objetiva aprofundar o tema do estabelecimento da norma-padrão, dando continuidade à proposta segundo a qual os parâmetros da norma de referência, que deveriam refletir os das normas praticadas em diversas variedades/modalidades, precisariam admitir natural flexibilidade e adaptabilidade por contexto sociocomunicativo, sendo igualmente plurais. Partindo dessa premissa, as seções do texto (i) sistematizam os pressupostos para que as orientações normativas escolares sejam produtivas (Seção 1) e (ii) apresentam as bases para o funcionamento de tais orientações, valendo-se de trabalhos que descrevem o comportamento de variedades cultas quanto a dois fenômenos morfossintáticos em diversos gêneros textuais (Seção 2) e, ainda, em experiências pedagógicas para a abordagem de regras variáveis (Seção 3). Por fim, apresenta discussão (Seção 4) acerca das contribuições da proposta, formulando as situações para a prática de orientações normativas, e sublinha, a título de considerações finais (Seção 5), a pertinência da relação entre orientação normativa e conhecimento sociolinguisticamente fundamentado.

## INTRODUÇÃO/CONTEXTUALIZAÇÃO

Na agenda dos trabalhos do eixo relacionado à variação e ensino no âmbito do GT de Sociolinguística da ANPOLL, foram propostas, no encontro do grupo em 2017, três frentes de trabalho a serem perseguidas pelos pesquisadores da área: (i) sistematização das diversas regras variáveis em função de contínuos de variação (fala-escrita e monitoração estilística); (ii) investigação de crenças e atitudes em relação às variantes linguísticas e ao ideário de norma-padrão; e (iii) desenvolvimento de metodologias para o trabalho didático com normas e variação linguística. Considerando diversas pesquisas brasileiras acerca desses tópicos, pode-se observar que o (velho) tema da norma-padrão a ser veiculada nas orientações escolares figura sempre atual – e, de fato, nunca se esgota –, por seu próprio estatuto fundamental: o de servir de referência reguladora/padronizadora dos usos linguísticos em meio à complexa rede de variedades e estilos.

Partindo de alguns resultados científicos atinentes sobretudo a duas das referidas frentes de trabalho (a primeira e a terceira), o presente texto objetiva aprofundar o tema do estabelecimento da norma-padrão, dando continuidade à proposta elaborada em Vieira (no prelo), segundo a qual os parâmetros da norma de referência, que deveriam refletir os das normas praticadas em diversas variedades/modalidades, precisariam admitir natural flexibilidade e adaptabilidade, assumindo, em alguma medida, esse caráter plural. Com base em experiências pedagógicas fundamentadas no conhecimento científico, a premissa básica da proposta foi assim sintetizada:

(...) é possível conjugar orientações normativas e pluralidade de normas de uso, deixando nítido aquilo que *a priori* deveria ser óbvio: o postulado da norma-padrão – não obstante pressuponha escolhas por contexto – não é incompatível com heterogeneidade nos usos sociais (VIEIRA, no prelo).

Partindo dessa premissa, as seções seguintes deste texto (i) sistematizam e ampliam os pressupostos adotados por Vieira (no prelo) para que as orientações normativas escolares sejam produtivas (Seção 1) e (ii) apresentam as bases para o funcionamento de tais orientações valendo-se de trabalhos que descrevem o comportamento de variedades cultas quanto a dois fenômenos morfossintáticos em diversos gêneros textuais (Seção 2) e, ainda, em experiências pedagógicas para a abordagem de regras variáveis (Seção 3). Por fim, apresenta discussão (Seção 4) acerca das contribuições da proposta, formulando possíveis situações para a prática de orientações normativas, e sublinha, a

título de considerações finais (Seção 5), a pertinência da relação entre orientação normativa e conhecimento sociolinguisticamente fundamentado.

## PRESSUPOSTOS PARA ORIENTAÇÕES NORMATIVAS ESCOLARES

A fim de fundamentar as orientações normativas escolares, toma-se por base o amplo e fecundo debate acerca da polissemia da palavra *norma*, considerando diversas referências acerca do tema – Faraco 2008; Zilles; Faraco 2015; Faraco; Zilles 2017; Görski; Coelho 2006; 2009; Vieira; Brandão 2007; Lagares; Bagno 2011; Bagno 2012; Görski; Freitag 2013; Duarte 2013; Kato 2013; Vieira 2013; Cyranka 2014; Martins; Vieira; Tavares 2014; Duarte; Serra 2015; Martins 2017; Freitag 2017; Barbosa; Vieira 2017; Vieira 2018a; Vieira 2018b; Vieira 2019a, Vieira 2019b, dentre outros. A partir desse conjunto de reflexões, a autora ratifica a necessária relação entre dois campos significativos relacionados ao termo *norma*, o da norma-padrão (*norma normativa*) e o das normas de uso (*normas normais*), dentre os quais figuram as variantes cultas.

Não obstante a complexidade que se esconde por trás da oposição entre *norma-padrão* e *norma culta*, vistas como blocos conceituais distintos, assume-se, aqui, em caráter tão-somente didático, o que seria prototípico de cada polo dessa oposição. Por *norma-padrão*, entende-se aquela que funcionaria como referência reguladora/uniformizadora das escolhas do que dizer/escrever em cada contexto sociointeracional, constituindo-se como instância do plano da idealização subjetiva e da prescrição. Segundo Faraco (2008), trata-se de “uma codificação relativamente abstrata, uma baliza extraída do uso real para servir de referência, em sociedades marcadas por acentuada dialeção, a projetos políticos de uniformização linguística” (p. 75). Por *norma(s) culta(s)*, concebem-se efetivas variedades que reúnem o conjunto de usos linguísticos partilhados e avaliados por uma comunidade de fala, constituindo-se como instância do plano da realização objetiva e passível de descrição científica. Como delimitação do perfil dos falantes supostamente cultos, essa comunidade seria constituída sobretudo por indivíduos de alto nível de escolaridade e que têm acesso a experiências diversas da cultura letrada/escrita. Consoante Faraco (2008), trata-se do “conjunto de fenômenos linguísticos que ocorrem habitualmente no uso dos falantes letrados em situações mais monitoradas de fala e escrita” (p. 73).

A desejável relação entre *norma-padrão* e *norma culta* sustenta igualmente uma das premissas assumidas por Faraco (2015), ao tratar do desafio de propor orientações de natureza normativa para o contexto escolar: é necessário que “os

instrumentos normativos reflitam a norma efetivamente praticada” (p. 15). Enfrentar esse desafio implica romper o paradigma que sustenta alguns discursos normativos brasileiros, que, com o legítimo propósito de uniformização, propõem uma generalidade na aplicação de suas regras de comportamento social, para toda e qualquer situação comunicativa. Assim, acabam por anular – tácita ou assumidamente – a variabilidade natural e inerente aos usos linguísticos, cultos ou não, que se manifestam nas relações sociointeracionais. Em outras palavras, há que se evitar uma falsa correlação entre a homogeneização proposta a cada aplicação de uma norma-padrão e a assunção de invariabilidade linguística, numa perigosa equação do tipo “norma-padrão = Língua Portuguesa”. O efeito desses discursos é o que justifica que alunos brasileiros cheguem, após alguns anos na escola, à terrível conclusão de que “sua língua materna é a língua mais difícil do mundo”.

Embora há muito se assuma que a *norma-padrão* deva se atualizar em função dos usos linguísticos efetivamente praticados por falantes considerados cultos – e não simplesmente reproduzir as regras prescritas em manuais tradicionais elaborados segundo modelos que perduram secularmente, sob pena de ser caduca e estéril –, uma questão ainda parece sem resposta objetiva: como tornar concreta a premissa de que a norma-padrão reflita, ao menos, a(s) norma(s) culta(s), no contexto pedagógico? A dificuldade parece residir na própria concepção de padronização: padronizar implica admitir a pluralidade e uma escolha específica em meio a essa pluralidade; trata-se necessariamente do polo da unidade (e conseqüente homogeneidade/invariabilidade) em contraposição ao polo da variabilidade. Como, então, conjugar ou adaptar a suposta unidade de um padrão a uma realidade que é naturalmente variável e heterogênea?

Aquilo que é teoricamente claro e de fácil assunção não parece ecoar com a mesma fluidez no campo aplicado. É urgente colaborar, então, para que os profissionais de ensino disponham de informações sobre as normas efetivamente praticadas que supostamente deveriam guiar as orientações normativas. Dispondo dessas informações, é inevitável assumir que, em diversos fenômenos linguísticos, as variedades cultas fazem uso das formas alternantes em função de modalidade (do maior ou menor grau de oralidade/fala-letramento/escrita) e registro (do mais ao menos formal/monitorado). Desse modo, a conseqüência natural para a orientação normativa escolar é a de que a padronização precisa ser efetivamente sensível a cada espaço desses contínuos.

Em termos práticos, deve-se admitir que, a cada instância de fala e de escrita, mais ou menos monitorada, o usuário da língua pode aplicar padrões com

particularidades. A título de exemplificação, como um falante culto se comporta em relação à expressão de acusativo anafórico de terceira pessoa (*encontrei-o* / *encontrei ele* / *encontrei Ø* / *encontrei o rapaz*)? Farta literatura sociolinguística brasileira (DUARTE; RAMOS 2015; VIEIRA; FREIRE 2014) permite afirmar que a estratégia *encontrei Ø* é opção preferencial na fala culta urbana e apresenta expressivo uso na escrita de textos menos controlados em termos estilísticos. Em textos de alta monitoração, a forma clítica passa a ser, entretanto, a opção preferencial, comutando com a estratégia com SN e com a forma vazia<sup>1</sup>.

Cabe, então, a reflexão: em função desse contínuo compósito fala-escrita/monitoração aqui brevemente ilustrado, caberia afirmar que a forma clítica (*encontrei-o*) é a única variante que deve estar representada e aceita na norma-padrão escolar, de modo uniforme? Uma norma-padrão uniforme a esse ponto não só impediria a premissa de que “os instrumentos normativos devem refletir as normas praticadas”, mas seria altamente improdutiva em relação às orientações curriculares nacionais, que pressupõem o trabalho com a língua socialmente situada segundo a expressão de diversos gêneros textuais da fala e da escrita.

Em defesa de uma norma-padrão plural – no sentido de sensível aos espaços do contínuo compósito fala-escrita mais ou menos monitorada –, é preciso avançar em ações para a construção de uma “pedagogia da variação linguística” (conforme identifica FARACO, 2008). Para tanto, duas reflexões adicionais devem estar na agenda dos estudos variacionistas aplicados ao contexto educacional: (i) Todas as regras variáveis praticadas nas instâncias de fala e de escrita brasileiras já têm seu comportamento descrito e/ou sistematizado para que possam servir de base aos postulados dessa norma-padrão?; e (ii) Já se experimentaram propostas pedagógicas que visem a cumprir esse intento?

Vieira e equipe de orientandos têm investido na resolução desse conjunto de questões que continuam atuais. A título de sistematização da proposta elaborada, cabe resumir e reiterar, aqui, alguns pressupostos assumidos pela autora em produções anteriores:

- (i) na esfera sociolinguística, constitui objetivo do ensino de Língua Portuguesa tornar o aluno capaz de reconhecer e/ou produzir o maior número possível de variantes linguísticas da fala e da escrita (e não apenas a re-

---

<sup>1</sup> Freire (2000; 2005) demonstra que o clítico acusativo de terceira pessoa figura em amostras de fala culta em apenas 3% das ocorrências, enquanto na escrita de textos menos monitorados/mais autorais, como crônicas jornalísticas e entrevistas impressas, passa a 44%, chegando a 73% em reportagens e editoriais, por exemplo.

afirmação da norma vernacular), promovendo a ampliação do repertório linguístico dos alunos;

- (ii) o trabalho escolar com formas alternantes da língua deve ser desenvolvido não só para o plano da produção textual, mas também para o da recepção (leitura/escuta);
- (iii) para tanto, devem ser priorizadas, na agenda de pesquisa na área conhecida como Sociolinguística Educacional (BORTONI-RICARDO, 2005), duas frentes de trabalho: (a) descrever, sistematizar e divulgar o comportamento e a avaliação de regras variáveis no contínuo fala-escrita, em diversos gêneros textuais, e no contínuo de monitoração estilística, em diversas situações sociointeracionais; e (b) desenvolver, experimentar e avaliar atividades didáticas que promovam o reconhecimento e/ou a produção de variantes linguísticas.

Em relação aos dois primeiros pressupostos, é essencial admitir que o aluno pode desenvolver autonomia para fazer escolhas conscientes quanto ao uso de variantes, o que é didaticamente facilitado pelo trabalho com o maior número possível de expressões faladas e escritas, apresentáveis concretamente nas instâncias de circulação social da linguagem, na forma de gêneros textuais. Nesse sentido, conforme explicita Vieira (no prelo), algumas questões auxiliam na reflexão:

- (i) Quem deve decidir se vai produzir/adotar determinada variante (conhecimento ativo) ou se vai apenas reconhecê-la (conhecimento passivo)? O próprio usuário da língua.
- (ii) Quem aceita ou rejeita as variantes que o usuário da língua escolhe? O interlocutor e, no contexto escolar, figura em destaque como interlocutor o professor, que, via de regra, avaliará as variantes que o aluno escolhe.
- (iii) Qual critério deve ser usado para a aceitação ou a rejeição das variantes? Ele deve estar fundamentado em resultados das pesquisas científicas, os quais devem oferecer o mapeamento das regras variáveis em função dos contínuos de urbanidade, oralidade-letramento e monitoração estilística.

No que se refere às frentes de trabalho a serem perseguidas na agenda da Sociolinguística Educacional Brasileira que, em última instância, poderão colaborar para o desenvolvimento da referida autonomia discente em um efetivo letramento sociolinguístico, as seções seguintes apresentam algumas iniciativas para o cumprimento de dois propósitos: (i) o de sistematização do comportamento das regras variáveis, sobretudo, no contínuo fala-escrita; e (ii) o de proposição de natureza metodológica para o trabalho pedagógico com regras variáveis.

## SISTEMATIZAÇÃO DO COMPORTAMENTO DAS REGRAS VARIÁVEIS NO CONTÍNUO FALA-ESCRITA

Não seria tarefa fácil elencar as diversas iniciativas brasileiras no sentido de colaborar com a meta de dar conta das regras variáveis consoante os muitos e variados gêneros textuais. Restringindo-se ao âmbito da Sociolinguística, duas obras produzidas pela equipe do GT da ANPOLL são emblemáticas no esforço de oferecer um mapeamento de variedades brasileiras quanto a alguns fenômenos: em Martins; Abraçado (2014), resultados científicos de diversas investigações acerca de temas fonéticos (vocalismo, consonantismo e entoação) e morfossintáticos (quadro pronominal – *tu versus você; nós versus a gente*; acusativo, dativo e reflexivo) foram sistematizados considerando, sobretudo, a fala brasileira; em Vieira; Freire (2014), na obra organizada por Martins; Vieira; Tavares (2014), o comportamento de cinco fenômenos morfossintáticos (concordância verbal, indeterminação do sujeito, clíticos acusativo e dativo, ordem dos pronomes átonos, alternância *ter versus haver* existenciais) em produções de indivíduos com curso superior (variedades cultas) foi sistematizado em função de instâncias do contínuo fala-escrita e monitoração estilística.

Conforme defende Vieira (no prelo), “produções dessa natureza contribuem para que o profissional de ensino ateste o que, cientificamente, parece óbvio: a norma culta, como qualquer variedade, é variável”. Assim, a descrição científica das regras variáveis pode sustentar a orientação relativa à chamada adequação das variantes a cada gênero textual, deixando de resultar das impressões e preferências estilísticas de revisores e professores de Português. Outro efeito importante da sistematização do comportamento das regras variáveis é o que se refere a demonstrar que não se justifica a manutenção de exigências quanto ao uso de formas linguísticas improdutivas ou até inexistentes no Português do Brasil, visto terem sido substituídas por outras em processos naturais de mudança ou configurarem tradições discursivas exclusivas de alguns gêneros textuais. Manter exigências de produção dessas formas em exercícios/treinos escolares acaba por promover o que acertadamente Faraco (2008) identifica como uma *norma curta*.<sup>2</sup> No que se refere a esses usos linguísticos, as orientações escolares

---

<sup>2</sup> Segundo Faraco (2008), a norma curta constitui um “conjunto de preceitos dogmáticos que não encontram respaldo nem nos fatos, nem nos bons instrumentos normativos, mas que sustentam uma nociva cultura do erro e tem impedido um estudo adequado da nossa norma culta/comum/standard” (p. 92).

poderiam ficar limitadas ao exercício do reconhecimento e da compreensão de tais estruturas em atividades de leitura/escuta.

Embora as referências citadas demonstrem o esforço em colaborar com a descrição das variedades brasileiras, tanto no contínuo fala-escrita, quanto no da monitoração estilística, entende-se que é preciso desenvolver essa descrição de forma orgânica e planejada. Buscando desenvolver projeto que dê conta desse propósito, iniciou-se – no primeiro semestre de 2018, como atividade em disciplina da Pós-Graduação em Letras Vernáculas, da Faculdade de Letras da UFRJ<sup>3</sup> – a composição de um banco de textos distribuídos segundo gêneros em um suposto contínuo de oralidade/letramento e monitoração estilística. Essa iniciativa deu origem aos primeiros empreendimentos do projeto científico intitulado *Contínuos de/em variedades do Português: análises contrastivas*, que tem por um de seus principais objetivos descrever, em um mesmo banco de dados e com os mesmos procedimentos metodológicos, o comportamento de regras variáveis em uma diversidade de gêneros textuais, da fala e da escrita.

Ainda na esfera dos trabalhos desenvolvidos na referida disciplina, foi possível levantar e sistematizar os dados referentes a oito fenômenos morfossintáticos em diversas instâncias de variedades cultas, quais sejam: estratégias de indeterminação do sujeito (clítico *se*, formas pronominais, expressões nominais); preenchimento do objeto direto (clítico, zero, SN, pronome reto); preenchimento do objeto indireto/dativo (clítico *lhe(s)*, zero, *a/para ele, a* mais SN); colocação pronominal (próclise, ênclise, mesóclise); expressão de futuro (forma analítica, forma simples, presente do indicativo); construções existenciais (*ter versus haver*); orações relativas (cortadora, copiadora, padrão).

Dez gêneros textuais – sendo apenas o primeiro deles, a entrevista sociolinguística, de fonte oral, embora alguns outros, como entrevistas impressas, tirinhas e anúncios, por exemplo, sejam de registro escrito mas concepção oral

---

<sup>3</sup> Participaram da tarefa de constituição do banco de textos e elaboraram trabalhos finais os seguintes alunos da Pós-Graduação, que cursaram a disciplina intitulada Tópicos Especiais, ministrada por Silvia Rodrigues Vieira, no primeiro semestre de 2018: Adriana Cristina Lopes Gonçalves, Bruna Brasil Albuquerque de Carvalho; Camila Nunes de Melo; Cristiane Barbalho da Silva Gaio de Sá; Daniela Gonçalves Ribeiro da Silva; Deyse Edberg Ribeiro da Silva; Eneile Santos Saraiva; Joyce Coutinho Nóbrega de Araújo; Juliana Cristina Vasconcellos Garcia, Juliana Magalhães Catta Preta de Santana; Karen Cristina da Silva Pissurno; Luan de Sousa Guimarães; Luzia de Cássia Almeida Passos; Maitê Lopes de Almeida; Michael de Araujo Palmieri; Monique Débora Alves de Oliveira Lima; Pedro Henrique Regis dos Santos; Rachel de Carvalho Pinto Escobar Silvestre; Robson Borges Rua; Thaissa Frota Teixeira de Araujo Silva; Vanessa Antunes da Silva.

(MARCUSCHI 2008) – serviram de base à referida descrição: entrevistas socio-linguísticas<sup>4</sup>; tirinhas; anúncios; entrevistas impressas (em jornais ou revistas); cartas de leitor; crônicas jornalísticas; notícias; editoriais; teses/dissertações; artigos em revistas científicas também da área da Comunicação Social. Quanto às fontes para a composição do *corpus*, os gêneros de jornal e revista foram extraídos, em sua maioria, de veículos de circulação pública, em sua maioria destinados à classe média-alta e alta, como *O Globo*, *Folha de São Paulo*, *Jornal do Brasil* e, mais raramente, *Extra*. Os textos acadêmicos foram extraídos de bancos de tese e revistas acadêmicas da área de Comunicação Social, produzidos na Região Sudeste.

Embora se encontre em curso a organização de livro dedicado à divulgação detalhada dessas investigações, parte dos resultados desse empreendimento serviu como ilustração da proposta feita em artigos científicos enviados recentemente para divulgação ou já publicados.<sup>5</sup> Do mesmo modo, apenas para ilustrar o alcance e a relevância do mapeamento sistemático das regras variáveis em função do contínuo de gêneros textuais da fala-escrita, apresentam-se, a seguir, resultados de duas das regras variáveis investigadas, quais sejam:

- (i) Construção existencial: *ter* (*tem muitas coisas para discutir*) versus *haver* (*há muitas coisas para discutir*)
- (ii) Expressão de futuro: forma simples (*eu viajarei amanhã*) versus outras formas – perífrase *ir* mais infinitivo (*eu vou viajar amanhã*) e presente do indicativo (*eu viajo amanhã*)

Na Tabela 1 e no Gráfico 1, a seguir, estão representados os índices referentes à variante considerada de maior prestígio em cada caso, aquela que, a rigor, seria privilegiada como pertencente a uma norma-padrão uniforme, quais sejam: a forma *haver*, no caso das construções existenciais; e a forma simples de futuro, na expressão de futuridade.

---

<sup>4</sup> Conferir amostra do Rio de Janeiro do Banco de dados *Concordância* do Projeto *Estudo comparado dos padrões de concordância em variedades africanas, brasileiras e europeias do Português*. Disponível em [www.corporaport.lettras.ufrj.br](http://www.corporaport.lettras.ufrj.br).

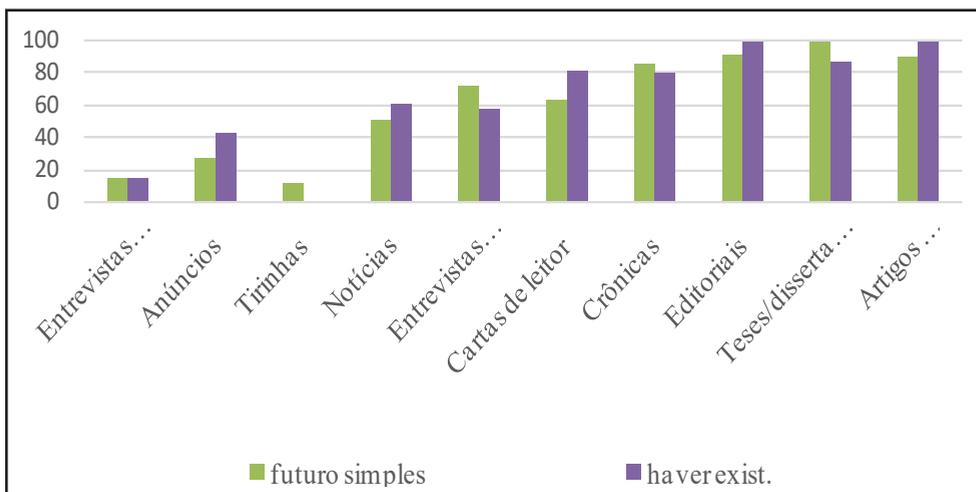
<sup>5</sup> Vieira (2018b), em artigo relacionado a unidade e diversidade no ensino de Língua Portuguesa, apresenta resultados relativos à alternância *ter* versus *haver* existenciais. Vieira (2019b), em artigo sobre o ensino de conectivos/conexão de orações, apresenta os resultados relativos às estratégias de relativização.

**Tabela 1:** Distribuição de dados de *haver* existencial e da forma simples de futuro em gêneros textuais da fala e da escrita cultas

Gênero textual	Haver existencial ( <i>versus ter</i> )	Futuro simples ( <i>versus</i> outras formas)
Entrevistas sociolinguísticas	107/725 (14,8%) <sup>6</sup>	5/35 (14%)
Tirinhas	2/2 (100%)	1/9 (11%)
Anúncios	3/7 (42,9%)	3/11 (27%)
Entrevistas em jornais e revistas	33/56 (58%)	48/67 (72%)
Carta do leitor	9/11 (81,8%)	22/35 (63%)
Crônicas Jornalísticas	17/21 (81%)	40/47 (85%)
Notícias	8/13 (61,5%)	14/28 (50%)
Editoriais	11/11 (100%)	30/33 (91%)
Teses/Dissertações	14/16 (87,5%)	27/27 (100%)
Artigos em revistas científicas	9/9 (100%)	18/20 (90%)

Fonte: Gama; Saraiva; Almeida (2018); Sá; Rua; Silva (2018)

**Gráfico 1:** Índices percentuais de *haver* existencial e da forma simples de futuro em gêneros da fala e da escrita cultas<sup>7</sup>



Fonte: Gama; Saraiva; Almeida (2018); Sá; Rua; Silva (2018)

<sup>6</sup> Apenas a coleta de dados de *ter/haver* relativa à fala (entrevistas sociolinguísticas) considerou a produção não só de indivíduos com Curso Superior mas também Fundamental/Médio.

<sup>7</sup> No caso da alternância *ter x haver* existenciais, não foi exibido no gráfico o comportamento dos dados extraídos das tirinhas, em função de terem sido registradas apenas duas ocorrências apenas da forma *haver*; entende-se que a representação desse índice categórico na imagem daria margem a uma interpretação generalizante infundada, dado o baixíssimo número de dados.

Os resultados referentes aos dois fenômenos representados permitem detectar o comportamento altamente variável da chamada norma culta: consoante o suposto grau de letramento ou de formalidade atribuídos ao gênero textual em questão, maior é o uso da variante considerada de prestígio. Embora os índices e sua escalaridade não sejam idênticos quanto aos dois fenômenos linguísticos, chama a atenção a semelhança das tendências encontradas. Em termos visuais, podemos reconhecer ao menos três porções/segmentos de realização das variantes representadas, quais sejam: os três primeiros gêneros apresentam baixos índices de realização das formas, enquanto os três últimos apresentam as mais altas taxas; na porção intermediária, os quatro gêneros restantes exibem índices medianos de realização das formas. Vieira (2018b), observando tão-somente a expressão de *haver versus ter*, propôs um contínuo que representasse esses segmentos dos gêneros em relação ao fenômeno. De forma semelhante, representa-se, na Figura 1, o comportamento dos dados em relação aos dois fatos variáveis em questão.

**Figura 1:** Contínuo de gêneros textuais: uso de *haver* existencial e forma simples de futuro

Entrevistas sociolinguísticas	Entrevistas impressas Cartas de leitor	Teses/dissertações Artigos científicos
Anúncios	Notícias	Editoriais
Tirinhas	Crônicas	
14-43%	58-85%	90-100% <sup>8</sup>

Embora os resultados sejam relativos a apenas dois fenômenos e considerando o fato de o banco de textos de onde se extraíram os dados lidar com uma pluralidade de suportes e condições interacionais, chama atenção a sistematicidade dos resultados a que se chegou, o que permite advogar em prol das acertadas motivações para a elaboração do referido projeto de investigação, ainda em fase inicial. Por ora, fica evidente o amplo espectro de variação inerente ao uso da forma prestigiosa em variedades cultas (que registra índices que vão de 14 a 100%) consoante a modalidade (mais oralidade – mais letramento) e a suposta monitoração estilística (mais ou menos controlada em função de diversas condições sociointeracionais).

<sup>8</sup> No extremo de letramento/escritura, deve-se salientar que a qualidade dos dados também é particular. Os poucos dados da forma “ter” registrados, por exemplo, também se diferenciavam, visto que os exemplos encontrados são da estrutura “tem-se”, em vez de simplesmente “tem”.

## DESENVOLVIMENTO DE METODOLOGIAS PARA O TRABALHO DIDÁTICO COM REGRAS VARIÁVEIS

Quanto à segunda frente de trabalho a ser priorizada pela Sociolinguística Educacional, a que se refere às investigações a partir de experiências didáticas para o tratamento de regras variáveis, já é considerável o número de pesquisas que se propõem a contribuir com a elaboração, experimentação e avaliação de metodologias para a construção de atividades pedagógicas.

No campo do ensino de regras variáveis, trabalhos recentes orientados com base na proposta de *Ensino de gramática em três eixos* (VIEIRA 2017a; 2017b; 2017c; 2018a) têm sido elaborados, sobretudo no âmbito do Mestrado Profissional em Letras, em rede nacional. A seguir, listam-se os temas já tratados em investigações a partir dessa abordagem; trata-se de pesquisas que, além de apresentarem resultados das variedades praticadas pelos alunos, propõem, aplicam e avaliam atividades para o ensino de fenômenos variáveis, quais sejam: a) indeterminação do sujeito (SOUZA 2014); expressão de segunda pessoa discursiva (GOUVÊA 2015); concordância verbal de terceira pessoa (CHAGAS 2015); quadro pronominal (LIMA 2016); expressão verbal de futuro (DURVAL 2017); colocação pronominal (ALMEIDA 2018).<sup>9</sup>

O tratamento dessa variedade de temas tem permitido testar a validade dos pressupostos assumidos quanto ao ensino de fenômenos variáveis. A esse respeito, fica claro que essas experiências permitem não só avaliar a distância entre as variedades praticadas pelos alunos e os usos esperados em gêneros de alto grau de letramento e monitoração estilística, mas também possibilita desenvolver a pluralidade de normas de uso, mesmo no âmbito da chamada norma culta. Para o cumprimento desses objetivos, as atividades<sup>10</sup> didáticas propostas costumam desenvolver:

- (i) atividades linguísticas: priorizar, em tarefas de leitura e produção textual, o contato intenso com textos de diferentes gêneros, em que as formas alternantes são naturalmente empregadas;

---

<sup>9</sup> O acesso total ao conjunto de atividades desenvolvidas depende da consulta às próprias obras mencionadas. Ademais, vale consultar artigos que apresentam sínteses de algumas dessas propostas pedagógicas (cf., por exemplo, VIEIRA 2018a; SOUZA; VIEIRA 2018; VIEIRA, 2019).

<sup>10</sup> A categorização de atividades aqui utilizada inspira-se em Franchi (2006), obra que publica os textos elaborados pelo autor na segunda metade do século XX e, conforme esclarece Sírio Possenti, na apresentação da obra, produtivamente utilizados na década de 1980.

- (ii) atividades epilinguísticas: construir – em etapas progressivas de aprendizagem conduzidas de forma indutiva, próprias de um estudo dirigido – o conhecimento necessário para a identificação das variantes, de seus efeitos de sentido e dos contextos linguísticos e extralinguísticos favoráveis a sua realização, além do valor social (se de estereótipo, marcador, indicador, nos termos de LABOV 1972) supostamente atribuído às variantes;
- (iii) atividades metalinguísticas: tornar consciente, por meio de um “trabalho inteligente de sistematização gramatical” (FRANCHI 2006, p. 98), o conhecimento adquirido por meio das reflexões e das conclusões a que as atividades conduziram em termos teóricos e descritivos.

Dessa forma, o trabalho com os fenômenos variáveis integra naturalmente um dos eixos do ensino de gramática, sendo esta também concebida como sistema que obedece naturalmente ao princípio da heterogeneidade ordenada. Embora os limites do presente texto não permitam descrever cada referida experiência pedagógica, essa frente de trabalho tem oferecido o necessário respaldo ao investimento que deve ser feito para a sistematização das regras variáveis em um contínuo de gêneros textuais, de um lado, e para a elaboração de metodologias para a chamada pedagogia da variação linguística, de outro.

## DISCUSSÃO

Consideradas as duas frentes de trabalho sinteticamente apresentadas, cabe avaliar as estratégias possíveis para o estabelecimento da chamada norma-padrão e refletir sobre seus limites em uma orientação pedagógica que se queira fundamentada no conhecimento das normas de uso. Em outras palavras, retomase o debate quanto à formulação de recursos e estratégias para tornar possível a premissa de que as orientações normativas devam tomar por base as variedades cultas em uso.

Em termos teórico-descritivos, fica evidente que a descrição das regras variáveis em uma diversidade de gêneros textuais de concepção ou fonte falada e escrita é absolutamente necessária. Os resultados relativos a apenas dois fenômenos apresentados no presente texto, a alternância *ter versus haver* existenciais e a expressão de futuro, são suficientes para demonstrar (i) a produtiva variabilidade interna à chamada norma culta e (ii) a consequente impossibilidade de propor uma norma-padrão única se atendida a premissa de que esta deva refletir a norma praticada. Em outras palavras, se a norma culta compreende um conjunto variável de formas (que vão de baixíssimos a

altíssimos índices a depender do contexto/gênero textual), seria absolutamente improdutivo propor orientações pedagógicas uniformes para todas as situações de expressões/instâncias cultas.

Quanto às propostas pedagógicas brevemente citadas, as pesquisas em questão asseguram que é possível dar a conhecer a extensão dos usos variáveis na diversidade de materiais e gêneros trabalhados, sem negar o valor social da linguagem e o prestígio/desprestígio das formas alternantes. Formular orientações normativas implica, portanto, sensibilidade em relação à tarefa em questão (Trata-se de que gênero textual? Da fala ou da escrita? Endereçado a que interlocutor? Em que suporte? Sob que condições?) e ao fenômeno em análise (Quais são as variantes que o compõem? Como elas se distribuem nos contínuos de variação? Alguma(s) das formas alternantes recebe(m) avaliação social explícita/consciente? De que tipo? Alguma(s) das formas alternantes está(ao) em desuso, configurando-se, no máximo, como uma tradição discursiva na contemporaneidade?).

Mediante a configuração que se faça das formas alternantes, as orientações emergirão naturalmente. Nesse sentido, Vieira (no prelo) formulou situações básicas para a definição de tais orientações, que aqui se retomam com adaptações e se aprofundam:

(1) Produtivo registro na fala e na escrita em diversos gêneros

Se houver produtivo registro de determinada variante nos diversos segmentos dos contínuos fala-escrita e monitoração estilística, praticamente não serão necessárias orientações escolares quanto à padronização. Isto porque, dada a circulação e a aceitação geral da forma em questão, trata-se, muito provavelmente, de um indicador, em termos labovianos.

Essa situação dispensa exemplificação, visto compreender muitas estruturas da língua posta em funcionamento, sem avaliação expressa quanto à padronização.

(2) Produtivo registro sobretudo na fala, particularmente em determinadas variedades

Em outro extremo, figura a situação em que determinada variante esteja intimamente relacionada à fala de determinadas comunidades, como, por exemplo, as representativas das chamadas variedades populares, configurando uma espécie de traço linguístico descontínuo, com reconhecível valor indexical (identificando pessoas não escolarizadas ou, muitas vezes, de forma preconceituosa, supostamente não inteligentes, por exemplo). Nessa situação, serão necessárias

orientações expressas sobre o valor social das variantes em termos de prestígio ou desprestígio. Isto porque, dada a especificidade na circulação e na avaliação social da forma em questão, trata-se muito provavelmente de um estereótipo, em termos labovianos.

Podem ilustrar essa situação as construções verbais e nominais sem a marca explícita de plural, sobretudo em estruturas muito salientes (alta saliência fônica, ordem direta SVO, como em *nós falava* ou *os avião saiu*, por exemplo).

### (3) Produtivo registro sobretudo na escrita e de gêneros mais monitorados

Se houver produtivo registro quase exclusivo na escrita (ou em falas altamente controladas) de gêneros mais monitorados e de menor circulação social, como os produzidos em meios profissionais e acadêmicos, serão necessárias orientações expressas para a produção (se o indivíduo precisar utilizar tais variantes, por razões profissionais ou por interesse em termos de variação estilística, ou se o indivíduo for submetido a uma avaliação formal, como a de redações em concursos) e para a recepção das variantes (em termos de leitura/interpretação de produções alheias). Isto porque, dada a circulação restrita a determinados meios e controlada a depender da situação, trata-se, muito provavelmente, de um marcador, em termos labovianos.

Exemplos de estruturas referentes a essa situação podem ser a estratégia de relativização padrão, o uso de *haver* existencial, a forma relativa *cujo* e variantes, clíticos acusativos e dativos de 3ª pessoa acusativos, ênclises e próclises pronominais segundo as regras tradicionais de colocação, dentre outras.

### (4) Ausência de registro na fala e na escrita contemporâneas ou registro apenas em gêneros de circulação em uma determinada comunidade de prática

Nesse caso, não havendo efetivo registro de determinada variante nem na fala nem na escrita contemporâneas ou, se houver, estiver limitado a comunidades de prática/tradições discursivas específicas, a conduta em termos pedagógicos também precisa ser particular. Assim, as orientações devem estar restritas ao campo da recepção dos dados (reconhecimento pela leitura/escuta – do domínio religioso ou literário, por exemplo), e não da produção, sob pena de promoção de arcaísmos. A produção da referida variante só será viável, na verdade, se estiver vinculada a uma necessidade de um grupo específico (religiosos, escritores, advogados, por exemplo) e não de uma comunidade de fala geral. Em outras palavras, está-se, possivelmente, diante de uma variante

antiga ou estrangeira ou, em situação específica, um marcador estilístico para a construção de uma *persona*<sup>11</sup>.

Trata-se, aqui, de um caso particular de formas linguísticas em flagrante desuso. Pode-se exemplificar essa situação com formas tidas como “arcaicas”, como os pronomes de segunda pessoa plural (*vós, vos, vosso*) ou as formas clíticas contratas (*mo, ta, ma, lho* dentre outras).

Esse conjunto de situações articulado ao conhecimento do perfil das regras variáveis da Língua Portuguesa, nos contínuos de fala-escrita e monitoração estilística, poderá contribuir valiosamente para que guias normativos efetivamente reflitam as normas praticadas na complexa rede de variedades, modalidades e registros.

## CONSIDERAÇÕES FINAIS

Como propõe Vieira (no prelo),

Admitir que o estabelecimento da norma-padrão (norma normativa) pode ser sensível às configurações das normas de uso (normas normais), das mais cultas às mais populares, pressupõe um novo olhar sobre o perfil do profissional de Língua Portuguesa, um profissional cuja prática se ancora nos resultados científicos e que colabora, ainda, para a produção de novos resultados.

É preciso, assim, investir na realização de pesquisas que, tratando das diversas regras variáveis, se dediquem às duas frentes de trabalho desenvolvidas no presente texto, para a construção de uma pedagogia da variação linguística. Em outras palavras, cabe dar continuidade às investigações que se ocupem não só do mapeamento dos fenômenos em gêneros textuais da fala e da escrita, consoante diversos graus de formalidade, mas também da construção das estratégias didáticas que promovam, a partir de atividades linguísticas, epilinguísticas e metalinguísticas (FRANCHI 2006), o conhecimento da língua como um sistema a um só tempo sistemático, interacional e, também, heterogêneo.

## REFERÊNCIAS

ALMEIDA, Sullaine Cristina Martins de. *Diagnose e análise de ordem dos clíticos pronominais em turmas de 9º ano: uma proposta interventiva*. 2018. Dissertação (Mestrado Profissional em Letras/PROFLETRAS) – Faculdade de Letras, UFRJ, Rio de Janeiro, 2018.

---

<sup>11</sup> Os chamados estudos de 2ª onda e de 3ª onda (ECKERT 2004) põem em evidência efeitos específicos de determinados usos por motivações de ordem estilística.

BAGNO, Marcos. *Gramática pedagógica do Português Brasileiro*. São Paulo: Parábola Editorial, 2012.

BARBOSA, Afrânio Gonçalves; VIEIRA, Silvia Rodrigues (Orgs.). *Revista Diadorim – Revista de Pós-Graduação em Letras Vernáculas/UFRJ*, v. 19, n. 2. Rio de Janeiro: Faculdade de Letras / UFRJ, 2017.

BORTONI-RICARDO, Stella Maris. *Nós chegemos na escola, e agora? Sociolinguística e Educação*. São Paulo: Parábola Editorial, 2005.

CHAGAS, Danieli Silva. *Concordância verbal de terceira pessoa: descrição sociolinguística e proposta pedagógica em turmas do ensino fundamental*. 2016. Dissertação (Mestrado Profissional em Letras/PROFLETRAS). Faculdade de Letras, UFRJ, Rio de Janeiro, 2016.

CHAGAS, Danieli Silva. Concordância verbal: estratégias para o trabalho com os três eixos para o ensino de gramática. In: VIEIRA, Silvia Rodrigues (Org.) *Gramática, variação e ensino: diagnose & propostas pedagógicas*. 2. ed. rev. e ampliada. São Paulo: Blucher, 2018. p. 61-93.

CYRANKA, Lúcia. Avaliação das variantes: atitudes e crenças em sala de aula. In: MARTINS, Marco Antonio; VIEIRA, Silvia Rodrigues; TAVARES, Maria Alice (Orgs.) *Ensino de Português e Sociolinguística*. São Paulo: Contexto, 2014. p. 133-156.

DUARTE, Maria Eugenia Lammoglia. O papel da Sociolinguística no (re)conhecimento do Português Brasileiro e suas implicações para o ensino. *Revista Letra*. Ano VIII, v. 1 e 2, p. 15-30, 2013.

DUARTE, Maria Eugenia Lammoglia; RAMOS, Jânia M. Variação nas funções acusativa, dativa e reflexiva. In: MARTINS, Marco Antonio; ABRAÇADO, Jussara (Orgs.). *Mapeamento sociolinguístico do Português Brasileiro*. São Paulo: Contexto, 2014. p. 173-195.

DUARTE, Maria Eugenia Lammoglia; SERRA, Carolina. Gramática(s), ensino de Português e “adequação linguística”. *Matraga*, v.22, n. 36, p. 31-55, 2015.

DURVAL, Luiz Felipe da Silva. Uma experiência didática com o futuro do presente: reflexão linguística, variação e ensino. In: VIEIRA, Silvia Rodrigues (Org.) *Gramática, variação e ensino: diagnose e propostas pedagógicas*. 2. ed. rev. e ampliada. São Paulo: Blucher, 2018. p. 155-171.

ECKERT, Penelope. *Linguistic variation as social practice*. Oxford: Blackwell Publishing, 2004.

FARACO, Carlos Alberto. *Norma culta brasileira – desatando alguns nós*. São Paulo: Parábola Editorial, 2008.

FARACO, Carlos Alberto. Norma culta brasileira: construção e ensino. In: ZILLES, Ana Maria S.; FARACO, Carlos Alberto (Orgs.). *Pedagogia da variação linguística: língua, diversidade e ensino*. São Paulo: Parábola, 2015. p. 19-30.

FARACO, Carlos Alberto; ZILLES, Ana Maria S. *Para conhecer norma linguística*. São Paulo: Contexto, 2017.

FRANCHI, Carlos. *Mas o que é mesmo “gramática”?* São Paulo: Parábola Editorial, 2006.

FREIRE, Gilson Costa. *Os clíticos de terceira pessoa e as estratégias para sua substituição na fala culta brasileira e lusitana*. 2000. Dissertação (Mestrado em Letras Vernáculas). Faculdade de Letras, UFRJ, Rio de Janeiro, 2000.

FREIRE, Gilson Costa. *A realização do acusativo e do dativo anafóricos de terceira pessoa na escrita brasileira e lusitana*. 2005. Tese (Doutorado em Letras Vernáculas). Faculdade de Letras, UFRJ, Rio de Janeiro, 2005.

FREITAG, Raquel M. K. A mudança linguística, a gramática e a escola. *Percursos*, v. 18, n. 37, p. 63-91, 2017.

GAMA, Deyse Edberg Ribeiro Silva; SARAIVA, Eneile Santos; ALMEIDA; Maitê Lopes de. *Tem variação entre as formas verbais impessoais ter e haver nas modalidades oral e escrita, em realizações da norma culta, do português brasileiro?* 2018. Monografia da disciplina Tópicos especiais (Programa de Pós-Graduação em Letras Vernáculas) – Faculdade de Letras, UFRJ, Rio de Janeiro, 2018.

GERHARDT, Ana Flávia. *Ensino de gramática e desenvolvimento metalinguístico: teorias, reflexões e exercícios*. Campinas, SP: Pontes, 2016.

GÖRSKI, Edair M.; COELHO, Izete L. Variação linguística e ensino de gramática. *Working papers em Linguística*, 10 (1), p. 73-91, Florianópolis, 2009.

GORSKI, Edair M.; FREITAG, Raquel M. K. O papel da sociolinguística na formação dos professores de língua portuguesa com língua materna. In: MARTINS, Marco Antonio; TAVARES, Maria Alice (Orgs.). *Contribuições da Sociolinguística e da Linguística Histórica para o ensino de língua portuguesa*. Coleção Ciências da Linguagem Aplicadas ao Ensino, volume V. Natal: EDUFRN, 2013. p. 11-52.

GOUVÊA, Isabela P. *Variação das formas interlocutivas de segunda pessoa: estratégias pedagógicas*. 2016. Dissertação (Mestrado Profissional em Letras - PROFLETRAS). Faculdade de Letras, UFRJ, Rio de Janeiro, 2016.

KATO, Mary. A gramática do letrado: questões para a teoria gramatical. In: MARQUES, Mary et alii (Orgs.). *Ciências da linguagem: trinta anos de investigação e ensino*. Braga: CEHUM, 2005, p. 131-145.

LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.

LAGARES, Xoán Carlos; BAGNO, Marcos (Orgs.). *Políticas da norma e conflitos linguísticos*. São Paulo: Parábola Editorial, 2011.

LIMA, Monique Débora Alves de Oliveira. *Quadro de pronomes pessoais na escola: diagnose e proposta pedagógica* 2017. Dissertação (Mestrado em Letras Vernáculas). Faculdade de Letras, UFRJ, Rio de Janeiro, 2017.

MARCUSCHI, Luís. *Produção textual, análise de gêneros e compreensão*. São Paulo: Parábola, 2008.

MARTINS, Marco Antonio. Em defesa do ensino de gramática na escola. *Revista do GELNE*, v. 19, n. 1, p. 103-117, 2017.

MARTINS, Marco Antonio; ABRAÇADO, Jussara (Orgs.). *Mapeamento sociolinguístico do Português Brasileiro*. São Paulo: Contexto, 2014.

MARTINS, Marco Antonio; VIEIRA, Silvia Rodrigues; TAVARES, Maria Alice (Orgs.). *Ensino de Português e Sociolinguística*. São Paulo: Contexto, 2014.

SÁ, Cristiane Barbalho da S. G. de; RUA, Robson Borges; SILVA, Vanessa Antunes da. *Título da obra*. 2018. Monografia da disciplina Tópicos especiais (Programa de Pós-Graduação em Letras Vernáculas) – Faculdade de Letras, UFRJ, Rio de Janeiro, 2018.

SOUZA, Daniela da Silva. *Estratégias de indeterminação do sujeito: uma proposta pedagógica para o ensino de gramática*. 2015. Dissertação (Mestrado Profissional em Letras - PROFLETRAS). Faculdade de Letras, UFRJ, Rio de Janeiro, 2015.

VIEIRA, Silvia Rodrigues. Sociolinguística e ensino de português: para uma pedagogia da variação linguística. In: TAVARES, Marco Antonio; MARTINS, Maria Alice (Orgs.). *Contribuições da Sociolinguística e da Linguística Histórica para o ensino de língua portuguesa*. Coleção Ciências Aplicadas ao Ensino, volume V. Natal: EDUFRN, 2013. p. 53-90.

VIEIRA, Silvia Rodrigues. Três eixos para o ensino de gramática: uma proposta experimental. In: NORONHA, Claudianny Amorim; SÁ JR., Lucrécio

Araújo de. (Orgs.). *Escola, ensino e linguagem* [recurso eletrônico]. Natal-RN: EDUFRN, 2017a. <http://repositorio.ufrn.br>

VIEIRA, Silvia Rodrigues. Três eixos para o ensino de gramática. In: VIEIRA, Silvia Rodrigues (Org.). *Gramática, variação e ensino: diagnose & propostas pedagógicas*. Rio de Janeiro: Letras UFRJ, 2017b. p. 68-82.

VIEIRA, Silvia Rodrigues. Prática de análise linguística sem ensino de gramática? Reflexões e propostas. In: ATAÍDE, Cleber *et alii* (Orgs.). *Gelne 40 anos. Vivências teóricas e práticas nas pesquisas em Linguística e Literatura*. São Paulo: Blucher, 2017c. p. 299-318.

VIEIRA, Silvia Rodrigues (Org.). *Gramática, variação e ensino: diagnose & propostas pedagógicas*. 2. ed. rev. e ampliada. São Paulo: Blucher, 2018a.

VIEIRA, Silvia Rodrigues. A unidade e a diversidade no ensino de Língua Portuguesa. *Tabuleiro de Letras*, v. 12, n.3, p. 22-34, 2018b.

VIEIRA, Silvia Rodrigues. Objetivos pedagógicos e níveis gramaticais: um olhar sobre o ensino de Língua Portuguesa. In: PILATI, Eloisa; NAVES, Rozana; SALLES, Heloisa (Orgs.). *Novos olhares para a gramática na sala de aula: questões para estudantes, professores e pesquisadores*. Campinas, SP: Pontes Editores, 2019a. p. 67-91.

VIEIRA, Silvia Rodrigues. Ensinando gramática em três eixos: conectivos e conexão de orações. In: ROSÁRIO, Ivo *et alii* (Orgs.). *Anais do II Seminário do Grupo de Pesquisa Conectivos e Conexão de Orações*. v. 1, n.2. Niterói: Letras da UFF, 2019b. p. 57-79.

VIEIRA, Silvia Rodrigues. Contínuos de variação em sala de aula: o desafio de propor orientações normativas. In: BARONAS, Joyce *et alii*. (Orgs.). *Em torno (Entorno) da pedagogia da variação linguística*. Londrina: EDUEL, no prelo.

VIEIRA, Silvia Rodrigues; BRANDÃO, Silvia Figueiredo (Orgs.). *Ensino de gramática: descrição e uso*. São Paulo: Contexto, 2007.

VIEIRA, Silvia Rodrigues; FREIRE, Gilson Costa. Variação morfossintática e ensino de Português. In: MARTINS, Marco Antonio; VIEIRA, Silvia Rodrigues; TAVARES, Maria Alice (Orgs.). *Ensino de Português e Sociolinguística*. São Paulo: Contexto, 2014. p. 81-114.

ZILLES; Ana Maria S.; FARACO, Carlos Alberto. *Pedagogia da variação linguística: língua, diversidade e ensino*. São Paulo: Parábola, 2015.

# TOWARDS A FLEXIBLE STANDARD IN THE SCHOOL CONTEXT CONTRIBUTIONS FROM LINGUISTIC STUDIES

*Silvia Rodrigues Vieira (UFRJ)*

Considering the several Brazilian researches on the teaching of linguistic variation, the article aims to approach the establishing of the standard norm, assuming the proposal that its parameters should reflect those of the standards practiced in varieties/modalities. This proposal implies admitting natural flexibility and adaptability of the standard norm to the socio-communicative context. Based on this premise, the text systematizes and expands the assumptions that must be made for school normative guidelines to be productive, namely: (i) in the sociolinguistic sphere, it is a didactic aim to make the student capable of recognizing and/or producing as many linguistic variants (in speech and writing activities) as possible, promoting the expansion of the linguistic repertoire; (ii) alternating forms of language should be developed not only for the textual production scholar activities, but also for the reception ones (reading/listening); and (iii) in light of these assumptions, two priorities should be developed by Educational Sociolinguistics (BORTONI-RICARDO, 2005): (a) to describe, systematize and disseminate the use and evaluation of variable rules in the speech-writing *continuum*, in several textual genres, considering also stylistic textual profiles, in several socio-interactional situations; and (b) to develop, experiment and evaluate didactic activities that promote the recognition and/or production of language variants.

In theoretical-descriptive terms, it is evident that the description of variable rules in a variety of textual genres (of speech or writing) is absolutely necessary. The results concerning only two linguistic phenomena – the alternation of *ter* versus *haver* and the expression of future – are sufficient to demonstrate (i) the productive internal variability to the so-called educated varieties and (ii) the consequent impossibility to propose a single standard norm given the premise that it should reflect the norm practiced by educated people. In other words, if the educated norm comprises a variable set of forms (ranging from very low to very high indexes depending on the textual context/genre), it would be absolutely unproductive to propose uniform pedagogical guidelines for all situations of educated expressions/instances.

As for the Brazilian pedagogical proposals for the teaching of variable rules, the researches assure that it is possible to work with the extension of the variable uses in the diversity of textual genres, without denying the social value of the language and the prestige of certain forms. Such studies develop:

- (i) linguistic activities: in tasks of reading and textual production, they promote the intense contact with texts of different genres, in which the alternating forms are naturally employed;
- (ii) epilinguistic activities: in progressive stages of learning, conducted in an inductive way, they aim to construct the necessary knowledge for the identification of the variants, their semantic effects and the linguistic and extralinguistic contexts favorable to its realization, besides of social value (stereotype, marker, indicator, according to LABOV, 1972) supposedly attributed to the variants; and
- (iii) metalinguistic activities: by means of an “intelligent work of grammatical systematization” (FRANCHI, 2006, p. 98), they aim to make aware the knowledge acquired through the reflections and conclusions to which the activities lead in theoretical and descriptive terms.

Finally, the following basic situations for the definition of normative orientations are formulated:

1. If there is a productive use of a certain variant in several spoken and written texts (*indicator*), no school guidance on standardization is necessary.
2. At the other extreme, there is the situation in which a particular variant is closely related to the speech of certain communities, such as the so-called popular varieties. In this case, the variant is interpreted as a kind of

discontinuous linguistic trait with recognizable indexical value (*stereotype* of “not schooled” or often supposedly “not intelligent”, for example). In this situation, explicit guidance on the social value of variants in terms of prestige or discredit will be required.

3. If there is a productive use of a variant almost exclusively in written (or highly controlled speech) of more monitored genres (*marker*), such as those produced in professional and academic circles, explicit guidelines for textual production will be required (specially for professional reasons or for need of stylistic variation) and for the reception of the variants (in terms of reading / interpreting texts).

4. If there is no use of a particular variant in contemporary speech and writing (*archaism*) or if it is used only in a given community of practice, the conduct in pedagogical terms also needs to be particular. Thus, the guidelines should be restricted to the field of data reception (recognition by reading / listening, at religious or literary domain, for example). The production of this variant will only be feasible, in fact, if it is linked to a need of a specific group (religious, writers, lawyers, for example) or if it functions as a stylistic marker for the construction of a *persona* (according to ECKERT, 2006).

This set of situations – articulated to the knowledge of the variable rules of the Portuguese language, in the speech-writing and stylistic monitoring *continua* – may contribute to normative guides effectively reflect the norms practiced in the complex network of varieties, modalities and styles. It is necessary, therefore, to develop researches that, in dealing with variable rules, dedicate themselves to the construction of a “pedagogy of linguistic variation” (FARACO, 2008). In other words, it is necessary to deal with not only the mapping of phenomena in textual genres of speech and writing, according to different degrees of formality, but also the construction of didactic strategies that promote – from linguistic, epilinguistic and metalinguistic activities – the knowledge of language as a system at once systematic, interactional and also heterogeneous.

## REFERENCES

- BORTONI-RICARDO, Stella Maris. *Nós chegemu na escola, e agora? Socio-linguística e Educação*. São Paulo: Parábola Editorial, 2005.
- ECKERT, P. Communities of Practice. In: BROWN, K., (Ed.). *Encyclopedia of Language and Linguistics*, 2nd Edition, Elsevier, Amsterdam, 2006, p. 683-685.

FARACO, Carlos Alberto. *Norma culta brasileira – desatando alguns nós*. São Paulo: Parábola Editorial, 2008.

FRANCHI, Carlos. *Mas o que é mesmo “gramática”?* São Paulo: Parábola Editorial, 2006.

LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.

## ATITUDES LINGUÍSTICAS DE MESTRANDOS DAS UNIDADES DO PROFLETRAS DA REGIÃO DO TRIÂNGULO MINEIRO<sup>1</sup>

*Juliana Bertucci Barbosa (UFTM-Uberaba)*  
*Talita de Cássia Marine (UFU-Uberlândia)*

### RESUMO

Neste capítulo, buscamos apresentar resultados de um teste de atitudes linguísticas aplicado a professores da Educação Básica da rede pública de ensino. Tais professores são ingressantes do ano de 2018 no Programa de Mestrado Profissional em Letras (PROFLETRAS) de duas unidades da região do Triângulo Mineiro. Para construção do teste, baseamo-nos, principalmente, em fenômenos variáveis do Português Brasileiro, sendo um mais estigmatizado, como o fenômeno da concordância verbal, e outro menos estigmatizado, como os verbos “ter” e “haver” com sentido de existir. Acreditamos que compreender as atitudes linguísticas de professores de língua portuguesa pode subsidiar o planejamento de ações de conscientização linguística, formação continuada de professores e propostas de ensino do português.

### INTRODUÇÃO

Apesar de pesquisas como a de Labov (2008) sobre a variação fonética ocorrida no inglês falado na ilha de Martha’s Vineyard, na década de 1970, já

---

<sup>1</sup> Agências de fomento: Capes (PLI-Portugal) e CNPq (Edital Universal, Número do processo: 424520/2016-8).

apontarem para a importância de se investigarem as crenças, as atitudes e/ou a percepção linguística dos falantes, estudos dessa natureza ainda são recentes no Brasil, embora estejam em franca ascensão (CYRANKA, 2007, 2011; 2014; CYRANKA; RONCARATI, 2008; AGUILERA, 2008; BOTASSINI, 2009, CUBA; BARBOSA, 2013; CARDOSO, 2015, MARINE; BARBOSA, 2016, FREITAG et al., 2016, TEODORO, 2018, SENE, 2018, etc.). Na esteira desta tendência é que surge a presente pesquisa, partindo do pressuposto de que estudar as atitudes linguísticas de um grupo, por exemplo, pode auxiliar o pesquisador a compreender as reações dos falantes diante das diferentes variedades de uma língua ou, ainda, (re)conhecer o seu julgamento em relação a distintos usos linguísticos (GÓMEZ MOLINA, 1987, p.25).

Assim, neste capítulo, temos como objetivo principal apresentar resultados de um teste sobre as atitudes linguísticas aplicado a professores da Educação Básica da rede pública de ensino. Tais professores são ingressantes do ano de 2018 no Programa de Pós-Graduação Mestrado Profissional em Letras (Profletras), de duas unidades da região do Triângulo Mineiro. Para a construção do teste piloto, baseamo-nos, principalmente, em fenômenos variáveis do Português Brasileiro, sendo um mais estigmatizado – o fenômeno da concordância verbal (Cf. SCHERRE, 2005) – e, outro, menos estigmatizado – os verbos “ter” e “haver” com sentido de existir (Cf. SILVA, 2003).

## CONTEXTUALIZANDO E JUSTIFICANDO O ESTUDO COM ATITUDES

Por meio de investigações como a que ora apresentamos, podemos entender como as atitudes linguísticas interferem no processo de construção da identidade de uma comunidade, o que pode se constituir como fonte de pesquisa e de reflexão das mais diversas ordens. Isso porque, tal como afirma Almeida (2008, p. 277), acreditamos que “nossa maneira de falar, a linguagem que utilizamos para expressar nosso pensamento podem ser submetidos, também, a julgamentos de valor que lhe conferem lugar de prestígio ou desprestígio” no seio social, o que acaba por revelar a importância que os valores sociais refletem na linguagem, podendo condicionar e até determinar a aprovação ou a desaprovação de determinados usos. Essas avaliações, materializadas pelas atitudes linguísticas do falante diante de algumas variedades da língua ou diante de variantes específicas de determinadas variedades, segundo Moreno Fernández (1998, p. 179), podem agir decisivamente nos processos linguísticos de variação e mudança que se produzem nas comunidades de fala, visto que:

[...] una actitud favorable o positiva puede hacer que un cambio lingüístico se cumpla más rapidamente, que en ciertos contextos predomine el uso de una lengua en detrimento de otra, que la enseñanza-aprendizaje de una lengua extranjera sea más eficaz, que ciertas variantes lingüísticas se confinen a los contextos menos formales y otras predominen en los estilos cuidados. Una actitud desfavorable o negativa puede llevar al abandono y el olvido de una lengua o impedir la difusión de una variante o un cambio<sup>2</sup>.

Nesse sentido, orientadas pelos postulados dos psicólogos sociais Lambert e Lambert (1968), avaliamos, pautadas na aplicação de um questionário, como que professores de língua portuguesa, que atuam na Educação Básica, da rede pública de ensino e que são alunos ingressantes do curso de Pós-Graduação Mestrado Profissional em Letras (PROFLETRAS) de universidades públicas do Triângulo Mineiro, avaliam determinados fenômenos variáveis do Português Brasileiro, sendo um mais estigmatizado e, outro, menos estigmatizado.

Tal como Freitag e Santos (2016, p. 109), acreditamos que o que faz com que uma variável seja sensível ou não à avaliação de um determinado grupo de falantes, pode estar relacionado

ao seu grau de saliência, seja linguística, social ou ideológica. A observação não só do comportamento, mas das crenças e das atitudes em relação a determinado fenômeno linguístico, permite mensurar a percepção de uma variante de um fenômeno socio-linguístico.

Nesse sentido, acreditamos, também, que fenômenos linguísticos variáveis que são linguística e socialmente mais salientes, como a variação da concordância verbal, por exemplo, acabam estimulando atitudes linguísticas negativas, ao passo que fenômenos linguísticos menos salientes, seja do ponto de vista linguístico, seja do ponto de vista social –, por exemplo, o uso do verbo “ter” com sentido de “existir” –, por serem menos perceptíveis, acabam ocasionando atitudes linguísticas positivas (ou menos negativas).

(Re)conhecer essas atitudes linguísticas dos professores de língua portuguesa que atuam na Educação Básica, na rede pública de ensino, pode subsidiar o planejamento de ações de conscientização acerca da heterogeneidade

---

<sup>2</sup> “[...] uma atitude favorável ou positiva pode fazer que uma mudança linguística se cumpra mais rapidamente, que em certos contextos predomine o uso de uma língua em detrimento de outra, que o ensino-aprendizagem de uma língua estrangeira seja mais eficaz, que certas variantes linguísticas se confinem aos contextos menos formais e outras predominem nos estilos cuidadosos. Uma atitude desfavorável ou negativa pode levar ao abandono e ao esquecimento de uma língua ou impedir a difusão de uma variante ou uma mudança linguística” (tradução nossa).

linguística do Português Brasileiro, para além dos documentos oficiais que servem como diretrizes (Parâmetros Curriculares Nacionais-PCN) ou base (Base Nacional Curricular Comum-BNCC) ao ensino de língua portuguesa no Brasil, enfáticos ao defenderem um ensino de língua claramente envolto por uma perspectiva sociolinguística. Nesse sentido, é relevante ressaltar que a primeira dentre as dez competências específicas de língua portuguesa para o ensino fundamental que é apresentada pela Base Nacional Curricular Comum (BNCC, 2017, p. 85)<sup>3</sup> é a seguinte: “Compreender a língua como fenômeno cultural, histórico, social, variável, heterogêneo e sensível aos contextos de uso, reconhecendo-a como meio de construção de identidades de seus usuários e da comunidade a que pertencem”.

Além disso, conhecer e reconhecer as atitudes linguísticas dos professores de língua portuguesa em exercício na Educação Básica, da rede pública de ensino, pode contribuir sobremaneira para o planejamento de propostas de formação continuada desses professores, ainda muito carentes de formação sociolinguística com vistas ao ensino de língua portuguesa, tal como apontam pesquisas recentes, como as de Marine e Barbosa (2017), Teodoro (2018) e Sene (2018).

Cabe destacar que acreditamos, também, que nossa pesquisa pode nos dar pistas de como esses professores lidam com a variação linguística em sala de aula e como reagem frente a determinados fenômenos linguísticos variáveis, já que o julgamento sobre as variedades linguísticas que o falante emite, acaba refletindo, linguisticamente falando, na maneira como ele se comporta: como uma pessoa que tem “uma atitude mais normativa, mais purista” ou que tem “uma atitude mais tolerante” (CARDOSO, 2015, p. 10).

Além disso, há de se considerar que existem características extralinguísticas que podem influenciar na constituição da identidade linguística do falante, tais como idade, sexo, ocupação, grau de escolaridade, procedência (nação ou região de origem), entre outros, e que podem acabar contribuindo para a formação de critérios de valor que influenciam na avaliação que fazemos das pessoas e também das variedades linguísticas utilizadas por essas pessoas. Por isso, nosso questionário objetivou também, traçar o perfil social dos professores pesquisados<sup>4</sup>.

---

<sup>3</sup> A versão final da BNCC para o Ensino Fundamental foi homologada pelo Ministério da Educação (MEC), em 20 dezembro de 2017 e pode ser acessada por meio do seguinte link: <http://basenacionalcomum.mec.gov.br/wp-content/uploads/2018/02/bncc-20dez-site.pdf>.

<sup>4</sup> Na seção dedicada aos procedimentos metodológicos, descreveremos o questionário que elaboramos e aplicamos para esta pesquisa.

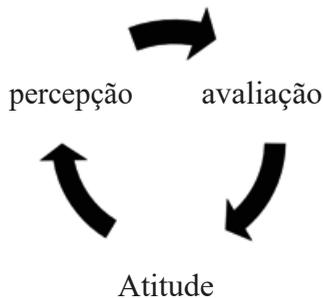
## PERCEPÇÃO, AVALIAÇÃO E ATITUDES LINGUÍSTICAS

Nesta pesquisa, assim como Freire (2016, p. 50), assumimos que a noção de percepção linguística:

[...] Está relacionada à capacidade cognitiva que os falantes têm em reconhecer diferenças linguísticas produzidas por eles mesmos e/ou por outros informantes em diferentes contextos das diversas esferas comunicativas. Assim, atribuem-se valores diferenciados ao que foi percebido (variante linguística com prestígio x variante linguística com desprestígio social, por exemplo). E essas práticas levam necessariamente os falantes a exibir atitudes sociolinguísticas.

Nesse sentido, a relação entre percepção, avaliação e atitudes linguísticas se configura por meio de uma atividade circular: o falante/ouvinte percebe diferenças de usos na língua e avalia essas diferenças, emitindo reações subjetivas acerca delas, manifestadas por meio de atitudes, tal como ilustrado por Freire (2016, p.51) pela figura abaixo:

**Figura 1:** Círculo permanente de práticas variacionistas



Fonte: FREIRE, 2016, p.51

Assim, parece-nos evidente que a noção de avaliação linguística está relacionada não apenas aos usos da língua, mas, também, aos seus usuários. Por isso,

semelhante à atitude linguística, torna-se uma prática sociolinguística, uma vez que não só são considerados aspectos relacionados à língua, mas também questões vinculadas ao próprio falante, seu lugar de origem, a outros fatores externos, como, por exemplo, sexo, idade, escolaridade, etnia, classe social, profissão, dentre outros (FREIRE, 2016, p. 51).

Já no que se refere à noção de atitudes linguísticas propriamente dita, Garrett (2010, p. 01) afirma que “*language attitudes permeate our daily lives. They are not always publicly articulated and, indeed, we are not always conscious of*

*them*”<sup>5</sup>. Nesse sentido, parece-nos evidente que as atitudes linguísticas, sejam elas positivas ou negativas, configuram-se como algo concreto e revelador (“o falante se manifesta”), pautadas nas mais diversas percepções, reações, avaliações e estereótipos que circulam numa dada comunidade de fala.

De acordo com Lambert e Lambert (1968), as atitudes se constituem de três componentes colocados no mesmo nível: o saber ou crença (componente cognoscitivo); a valoração (componente afetivo); e a conduta (componente conativo). Fazendo a transposição de tal conceito aos estudos sociolinguísticos, podemos afirmar que a dimensão comportamental (uso) se refere à produção, ou seja, como o falante efetivamente fala, a frequência de recorrência de uma dada variante em uma comunidade. Já as dimensões cognitiva e afetiva correspondem à percepção. Assim, de acordo com Freitag et al. (2016, p. 66):

Como o falante acha que fala ou acha que deve falar (cognitivo) é a manifestação verbalizada, sem reações afetivas, acerca da sua crença sobre seus usos e sobre os padrões da comunidade. Como o falante julga aqueles que falam de determinado jeito (afetivo) é a manifestação de reações afetivas em relação ao objeto em questão.

Assim, compreendemos a relação intrínseca entre percepção, avaliação e atitudes linguísticas e entendemos, tal como Garrett, Coupland e Williams (2003), que diversos são os termos que estão vinculados à noção de atitudes – hábitos, valores, crenças, opiniões e ideologias –, “o que significa dizer que a atitude linguística de um indivíduo é o resultado da soma de suas crenças, conhecimentos, afetos e tendências a comportar-se de uma forma determinada diante de uma língua ou de uma situação sociolinguística” (AGUILERA, 2008, p. 106). Logo, neste capítulo, focamos nossas reflexões nas atitudes linguísticas entendidas como tomadas de posicionamentos, julgamentos e reações, positivas ou negativas, a algo ou alguém, subdivididas em três dimensões: cognitiva (pensamentos e crenças), afetiva (sentimentos) e comportamental (uso).

## DESCRIÇÃO DOS PROCEDIMENTOS METODOLÓGICOS

Para a realização deste estudo, inicialmente aprovamos o projeto no Comitê de Ética em Pesquisa, via Plataforma Brasil. Como instrumento de pesquisa, elaboramos um teste de atitudes linguísticas<sup>6</sup> que foi aplicado a professores de Língua Portuguesa da Educação Básica, atuais alunos do Mestrado Profissional

<sup>5</sup> “Atitudes linguísticas permeiam nossas vidas diárias. Elas nem sempre são articuladas publicamente e, na verdade, nem sempre estamos conscientes delas” (tradução nossa).

<sup>6</sup> Alguns dos itens do teste partiram de resultados de Ghessi e Barbosa (2017).

em Letras<sup>7</sup> (Profletras) de duas universidades federais da região do Triângulo Mineiro: Universidade Federal de Uberlândia-UFU e Universidade Federal do Triângulo Mineiro-UFTM, campus sede em Uberaba.

## A ELABORAÇÃO E APLICAÇÃO DO TESTE DE ATITUDES LINGUÍSTICAS

Para esta pesquisa, realizamos um teste piloto de atitudes linguísticas para não só podermos refletir sobre algumas questões relacionadas à percepção e à atitude linguística, mas, também, para posteriormente podermos aprimorar o instrumento de coleta de dados. Para a montagem do instrumento de pesquisa – o teste de atitude –, consideramos, conforme afirmam Freitag et al (2016, p. 66, destaques nossos), que

além do tratamento societal, de caráter etnográfico, colhendo dados a partir de várias fontes de domínio público, como documentos oficiais, propagandas, televisão e, também, redes sociais (FREITAG et al., 2015), **as abordagens podem ser de forma direta** (perguntar aos falantes o que pensam sobre determinado fato da língua), mais propícia a captar a dimensão cognitiva, **ou indireta** (submeter os falantes à apreciação de características linguísticas e pedir que as associem a traços psicossociais atribuídos aos seus falantes, e, por tabela, à variante em questão), mais propícia a captar a dimensão afetiva.

Além disso, cabe destacar que, para medir a atitude, existem algumas metodologias já utilizadas em outras pesquisas, entre elas, o *matched guise test* e o *verbal guise test*. A técnica *matched guise test* busca identificar atitudes dos sujeitos em relação à língua (LAMBERT et al., 1960)<sup>8</sup>. Nesse teste, os participantes da pesquisa ouvem áudios de um mesmo falante e avaliam tais estímulos linguísticos, julgando-os quanto a aspectos afetivos (como, por exemplo, liderança, agradabilidade, aparência, questões sociais, etc). Outra técnica similar é o *verbal guise test*, que difere do *matched guise*, principalmente pelo fato de conter estímulos de diferentes falantes<sup>9</sup>. Por meio dessa técnica, segundo Freitag e Santos (2016, p.113), é possível “extrair ocorrências de variantes linguísticas de um fenômeno variável de um *corpus* oral previamente gravado, e utilizá-las na

<sup>7</sup> Os mestrandos são alunos ingressantes em março de 2018 nas duas universidades.

<sup>8</sup> Técnica desenvolvida por Lambert et al. (1960), foi originalmente usado para descobrir as atitudes de canadenses de Montreal para inglês e francês. Em seu estudo, apresentaram duas gravações, uma em inglês e outra em francês, aos alunos escola secundária bilingue. Tais alunos avaliaram as gravações, acreditando que cada uma delas era de uma pessoa diferente, embora, na realidade, fossem de um mesmo indivíduo, apenas em idiomas diferentes.

<sup>9</sup> Desta técnica, Agheyisi e Fishman (1970) desenvolveram uma técnica chamada *verbal guise test* em que os falantes das gravações são pessoas diferentes.

elaboração do protocolo de testagem”. Cabe mencionar que acreditamos que cada método tem vantagens e desvantagens; por isso, realizamos adaptações de acordo com o objetivo de nosso estudo, considerando, inclusive, as especificidades dos participantes da nossa pesquisa (mestrandos do Profletras que são também, obrigatoriamente, professores de língua portuguesa da Educação Básica).

Para a montagem do nosso teste piloto, baseamo-nos e adaptamos, principalmente, o *verbal guise test*, visto que utilizamos estímulos linguísticos de diferentes alunos da Educação Básica; entretanto, não partimos de texto oral, mas, sim, de fragmentos de produções escritas por alunos. Optamos por textos escritos para evitar interferências acústicas (como, por exemplo, julgamento do sotaque, que não era foco de nosso estudo) ou mudanças de padrão entoacional. Cabe ressaltar que, para selecionarmos os fragmentos que foram julgados pelos participantes da pesquisa, focalizamos em tais fragmentos os usos de dois fenômenos variáveis: um supostamente mais estigmatizado, a concordância verbal, e outro supostamente menos estigmatizado, variação no emprego de “ter” e “haver” com valor existencial.

A coleta dos dados foi realizada por meio de instrumento eletrônico, com o auxílio da ferramenta *Google Forms*<sup>10</sup>. O teste foi enviado para o e-mail dos informantes. Antes de responder ao teste, os participantes da pesquisa leram e aceitaram um termo de esclarecimento e consentimento que aparecia na primeira tela do formulário. Somente após o aceite de sua participação é que o respondente teve acesso ao teste. Nosso teste configura-se como um “piloto” de nossa pesquisa, pois, como já mencionado, a partir dos resultados obtidos, pretendemos refinar o teste que concebemos e aplicamos, o qual foi constituído de 16 itens, dividido em três partes:

**I Parte** – composto de perguntas sobre o perfil social e formação acadêmico-profissional dos mestrandos/respondentes do Profletras;

**II Parte** – para essa etapa do teste, apresentamos 04 fragmentos de textos escritos por alunos. Após cada fragmento, havia três questões. A primeira de múltipla escolha, que solicitava aos respondentes que avaliassem os fragmentos escolhendo adjetivos de diferentes campos semânticos – dia-tópico, aparência física, escolaridade, aparência intelectual, profissional, estilo de vida/social, aparência/agradabilidade, avaliação social –, e duas

---

<sup>10</sup> O *Google Forms* é uma ferramenta (gratuita para todos que possuem uma conta *Google*) que pode ser acessada em diversas plataformas, como web, desktop e celular e tem como objetivo principal a elaboração e aplicação de formulários de pesquisa ou de coleta de opiniões.

perguntas abertas, uma avaliativa (de julgamento em relação à escrita do aluno) e outra sobre tomada de atitude do professor, em sala de aula, em relação à escrita. Nessa última questão, nossa intenção não era identificar a atitude, mas, sim, observar se o professor mencionava (percebia) o fenômeno variável selecionado para nossa pesquisa.

**III Parte** – essa seção do teste está relacionada ao uso das variantes, pois solicitamos, a partir de 04 fragmentos de textos escritos por alunos, que o respondente indicasse em quais contextos ele orientaria um determinado aluno “a escrever e/ou a utilizar as formas linguísticas” dos fragmentos apresentados. Para isso, utilizamos uma escala de 1 a 5 (cf. CARDOSO, 2015), sendo 1 o contexto de escrita menos monitorado e 5 um contexto de maior monitoramento do falante. Abaixo temos um exemplo:

Figura 2: parte III do teste

:::					
<b>2) Há vários trabalhos para próxima semana.</b> *					
	1	2	3	4	5
Mensagem escrita a um amigo em rede social ou aplicativo de celular	<input type="radio"/>				
					Redação escolar/redação em vestibular
<b>3) Nós mudou porque a sala de cima era mais boa (M, 10, CB)</b> *					
	1	2	3	4	5
Mensagem escrita a um amigo em rede social ou aplicativo de celular	<input type="radio"/>				
					Redação escolar/redação em vestibular
<b>4) Nós jogamos futebol, vôlei, basquete e tênis de mesa. (M, 10, CB)</b> *					

Fonte: as autoras.

## **METODOLOGIA DE ANÁLISE DO TESTE**

Inicialmente, os dados obtidos após a aplicação do teste foram analisados separadamente: por universidade, por região, por escolaridade. Em seguida, como os resultados encontrados isoladamente, em sua maioria, foram similares, relacionamos e reunimos as informações.

Posteriormente, tabulamos os dados extraídos do questionário com auxílio do *Excel*, aplicativo *Windows*, e de ferramentas estatísticas do *Google Forms*, realizamos a análise exploratória, selecionando os componentes principais que poderiam nos ajudar a descrever o perfil (social/acadêmico/profissional) e as atitudes linguísticas dos mestrandos do Profletras das duas universidades da região do Triângulo Mineiro.

## **RESULTADOS**

Nesta seção, apresentamos os resultados da aplicação do teste de atitudes aos alunos do Profletras das duas universidades federais da região do Triângulo Mineiro: Universidade Federal de Uberlândia (UFU) e Universidade Federal do Triângulo Mineiro (UFTM-Uberaba). Cabe mencionar que, mesmo tendo verificado que em todas as respostas há resultados interessantes que revelam as atitudes linguísticas dos participantes da pesquisa, descrevemos e discutimos, neste capítulo, apenas alguns dados. Nosso critério de seleção levou em conta as respostas que se mostraram mais contundentes no que se referia à percepção e às atitudes dos professores, principalmente, em relação aos dois fenômenos variáveis escolhidos e à prática docente.

## **PERFIL SOCIAL E ACADÊMICO-PROFISSIONAL**

Ao enviarmos o teste por e-mail aos mestrandos ingressantes de 2018 do Profletras das duas universidades (37 alunos), obtivemos 16 respondentes, sendo 06 da UFTM e 08 da UFU.

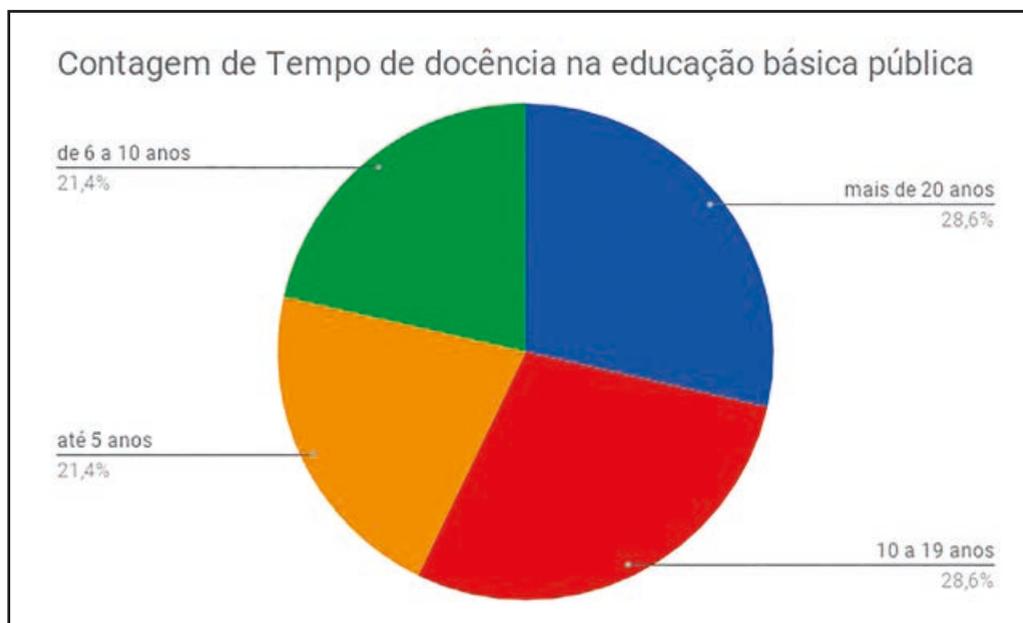
Em relação à região de origem dos mestrandos, assim como no trabalho sobre crenças linguísticas de Marine e Barbosa (2017), verificamos que o Profletras da UFTM possui alunos oriundos de Uberaba, de cidades da região do Triângulo Mineiro, do sul de Minas Gerais e do oeste do estado de São Paulo. Já o Profletras da UFU possui alunos de Uberlândia, de cidades da região do Triângulo Mineiro, de Brasília (cidades satélites) e de cidades do interior de Goiás.

Essa primeira parte do teste ainda apontou que tais mestrandos têm como perfil social e acadêmico-profissional as seguintes características:

- a) estão na faixa etária de 30 a 49 anos;
- b) são, predominantemente, do gênero feminino (93%);
- c) 57% deles fizeram o Curso de Letras em universidades/faculdades públicas e todos na modalidade presencial;
- d) a maioria (60%) terminou o Curso de Letras entre os anos 2000 e 2010, graduando-se, portanto, depois da publicação dos Parâmetros Curriculares Nacionais (PCN) e de muitas pesquisas que tematizam os problemas do ensino de língua portuguesa;
- e) 79% não cursaram outra graduação;
- f) 62% fizeram curso de pós-graduação *lato sensu* (geralmente, na área de ensino/metodologia de Língua Portuguesa ou supervisão de ensino), (6%) cursaram pós-graduação *lato sensu* na área de Educação Inclusiva e os demais possuem apenas graduação.

Em relação ao tempo de docência na Educação Básica pública, há professores com diferentes experiências, como podemos observar no gráfico abaixo:

**Gráfico 1:** tempo experiência na docência (no Ensino Fundamental)



Como pode ser observado no gráfico 1, há um certo equilíbrio em nossa amostra em relação ao tempo de docência na Educação Básica pública, o que, para o nosso estudo, pode ser um elemento bastante relevante para a análise geral de nossos dados, já que os resultados obtidos partem de professores que possuem diferentes experiências docentes.

## AS ATITUDES LINGUÍSTICAS DOS MESTRANDOS DO PROFLETRAS DA REGIÃO DO TRIÂNGULO MINEIRO

Como já mencionado, os testes sobre atitude linguística referem-se às partes II e III do questionário. Na parte II, apresentamos 04 fragmentos: dois que focalizavam o fenômeno da concordância verbal (CV) e dois que focalizavam as variantes “ter” e “haver” com sentido de “existir”.

Vejam os dois primeiros fragmentos<sup>11</sup>, escritos pelo mesmo aluno da Educação Básica, que visaram, principalmente, o fenômeno da CV, sendo o primeiro com marcações e o segundo com ausência de marcações de CV:

*Fragmento 1 - É claro que eles perceberam que muitas coisas são diferentes dos costumes que temos aqui no Brasil, mas, mesmo assim ficaram encantados com tudo [A06, F, 17, AB]*

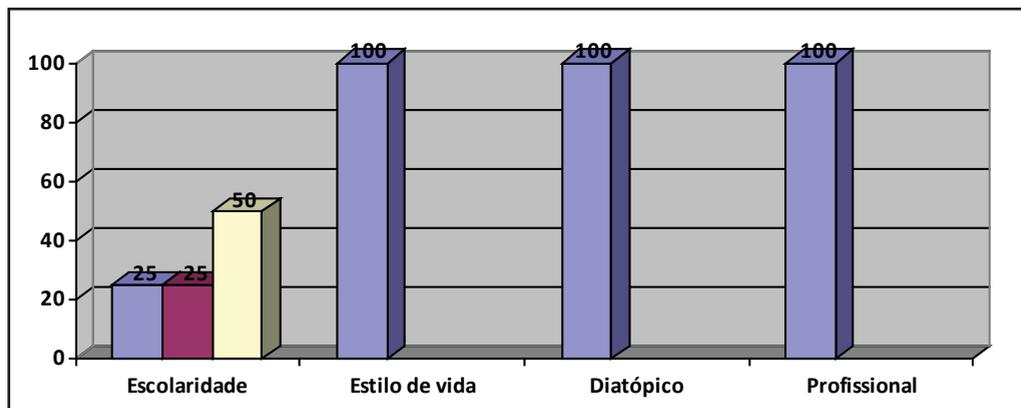
*Fragmento 2 – [...] eles ria e contava mais e mais situações que já lhe aconteres. [A06, F, 17, AB]*

Foi solicitado aos respondentes, a partir da leitura de cada um dos fragmentos, que selecionassem, a partir de caixas de múltipla seleção, expressões e/ou adjetivos de diferentes campos semânticos (diatópico, aparência física, escolaridade, aparência intelectual, profissional, estilo de vida/social, aparência/agradabilidade e avaliação social). Abaixo, apresentamos os itens selecionados – nem todos os itens foram escolhidos – pelos respondentes, acompanhados do percentual de ocorrência de tal seleção em nosso teste, a partir da avaliação dos fragmentos 1 e 2, respectivamente:

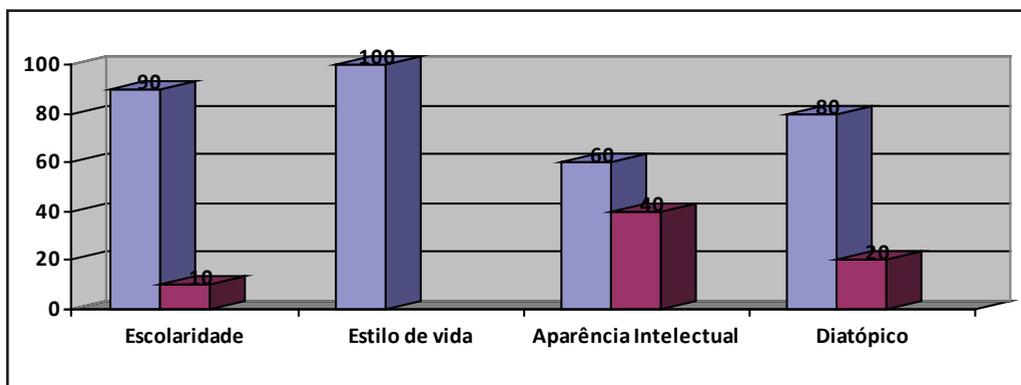
---

<sup>11</sup> Os fragmentos foram reproduzidos conforme texto original.

**Gráfico 2:** sobre Fragmento 1<sup>12</sup>



**Gráfico 3:** sobre Fragmento 2<sup>13</sup>



Como podemos observar nos gráficos 2 e 3, para o fenômeno da CV, alguns campos semânticos não foram selecionados pelos respondentes (são eles: aparência física, avaliação social) e apenas uma pessoa selecionou “Pessoa confiável” (critério de agradabilidade).

Em relação aos resultados da presença da CV (fragmento 1), a maioria dos respondentes julgou o fragmento 1 como escrito por uma pessoa com o seguinte perfil: Ensino Médio Completo (50%), estilo de vida “simples” (100%), pertencen-

<sup>12</sup> **Legenda:** **Escolaridade:** Ensino Médio Incompleto (25%), Ensino Fundamental (25%), Ensino Médio Completo (50%) / **Estilo de vida:** Simples (100%) / **Diatópico:** Zona Urbana (região central da cidade) (100%) / **Profissional:** Trabalhador empregado (100%).

<sup>13</sup> **Legenda:** **Escolaridade:** Ensino Fundamental Incompleto (90%), Ensino Médio Incompleto (10%) / **Estilo de vida:** simples (100%) / **Aparência Intelectual:** Pessoa com dificuldade de aprendizagem (60%), Não opinaram (40%) / **Diatópico:** Zona rural (80%) / Região Periférica da cidade (20%).

cente a zona urbana (região central) da cidade (100%), além de ser um “trabalhador empregado” (100%). Já em relação à ausência da CV, predominantemente, os respondentes avaliaram o fragmento 2 e atribuíram ao produtor daquele texto o seguinte perfil: ter Ensino Fundamental Incompleto (90%), ter um estilo de vida “simples” (100%), ser uma pessoa com dificuldade de aprendizagem (60%), ser da zona rural (80%) ou da região periférica da cidade (20%).

Tais resultados apontam que, para a ausência de concordância, os mestrandos do Profletras (professores da Educação Básica), ingressantes da turma de 2018, realizaram uma avaliação mais negativa, reforçando alguns estereótipos equivocadamente associados à ausência da CV, como sendo, por exemplo, um traço característico de uma pessoa oriunda da zona rural e com dificuldade de aprendizagem. Já para o fragmento 2, a avaliação se mostrou mais positiva, tal como podemos observar pela associação do referido fragmento como produção escrita de alguém pertencente à “zona central da cidade” e com maior grau de escolaridade.

Essa avaliação mais positiva para o fragmento 2 se comprova nas demais questões do teste, principalmente quando foi perguntado aos participantes da pesquisa, em relação aos fragmentos 1 e 2, se eles acreditavam que o autor de cada um dos textos “era um bom aluno”. Apenas um participante afirmou que não poderia avaliá-lo, tanto pelo fragmento 01, quanto pelo fragmento 02. Em relação ao fragmento 01, com presença de marcações de CV, a maioria dos respondentes avaliou da seguinte forma:

*(02) Sim, pois escreve bem com concordância, deixa claro a ideia e sem desvios ortográficos. (L11)*

*(03) Sim, pois escreve com clareza, é coerente e faz uso das concordâncias verbal e nominal. (L11)*

*(04) Sim, pois demonstra ter bastante conhecimento da língua portuguesa. (L11)*

Cabe mencionar que, pelos comentários feitos pela maioria dos respondentes, percebemos que houve, por parte desses professores/mestrandos, a avaliação de que o “aluno bom” é aquele que faz uso da concordância (verbal e nominal). O que, por oposição, pode revelar que tais professores acreditam (creem) que a não marcação da concordância pode estar associada ao “mau aluno”. Tal crença, por sua vez, pode impulsionar atitudes linguísticas em sala de aula bastante perniciosas e que podem, inclusive, contribuir para a baixa autoestima linguística

do alunado, impactando negativamente no desenvolvimento da competência comunicativa desses discentes.

Em relação ao fragmento 02, com ausência de concordância verbal, novamente apenas um respondente afirmou não poder avaliá-lo; os demais afirmaram que poderia sim “ser um bom aluno”, apenas com desvios de “norma padrão” ou com problemas de concordância, como em:

*(05) Sim, pois produziu um texto coerente porém com erros de concordância (L06)*

*(06) Precisa melhorar a concordância. (L08)*

Associando os comentários acima ao perfil social, acadêmico e profissional dos professores pesquisados, pelo qual sabemos que a maioria (60%) terminou o Curso de Letras entre os anos 2000 e 2010, graduando-se, portanto, depois da publicação dos Parâmetros Curriculares Nacionais (PCN), cuja orientação para um ensino de língua pautado nos gêneros textuais e sensível à variação linguística é evidente, acreditamos que essa avaliação positiva do aluno – com ressalvas do tipo “porém com erros de concordância” ou “precisa melhorar a concordância” – pode ser reveladora de uma avaliação “falsa positiva”. Nossa hipótese para justificar esse possível “falso positivo”, pauta-se na suposição de que tais professores, impulsionados pelo conhecimento das orientações dos PCN, bem como pelas inúmeras pesquisas publicadas no período em que cursavam a graduação – às quais pressupomos que tenham tido acesso durante a licenciatura –, que problematizaram a questão do ensino da gramática normativa na escola sem considerar o caráter heterogêneo da língua, tudo isso pode ter influenciado uma resposta “politicamente correta” ao teste, mas que, não necessariamente, representa o que de fato pensam sobre o aluno que não faz uso da concordância de acordo com as prescrições da gramática normativa. Nesse sentido, entendemos que uma nova questão deverá ser acrescida ao novo teste que conceberemos após a avaliação deste piloto, a fim de que possamos avaliar se nossa hipótese se confirma ou não.

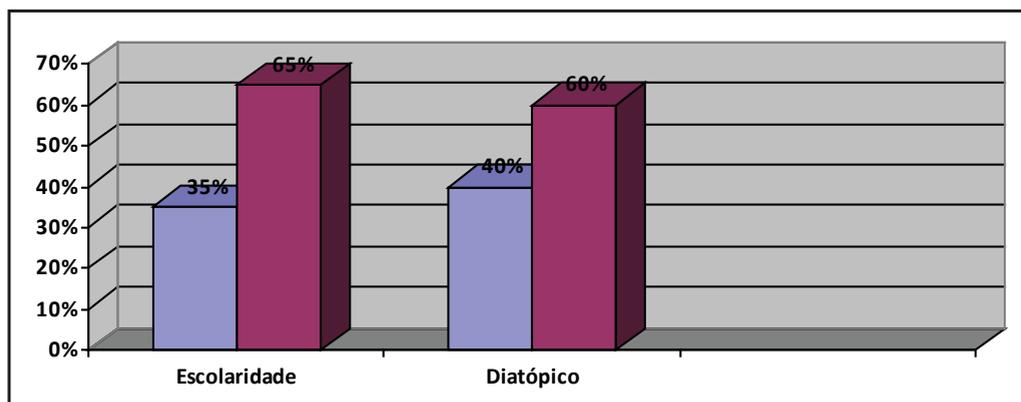
Vejam agora, o julgamento dos respondentes em relação a outros dois fragmentos (3 e 4), escritos por um mesmo aluno da Educação Básica. Os fragmentos foram selecionados para o teste devido ao emprego dos verbos “ter” e “haver” com sentido existencial:

*Fragmento 3 - Não tem essa história de pegar resto. Vem um monte de cara falar com ela, em alguns eu já estava de olho antes. [C82002P62]*

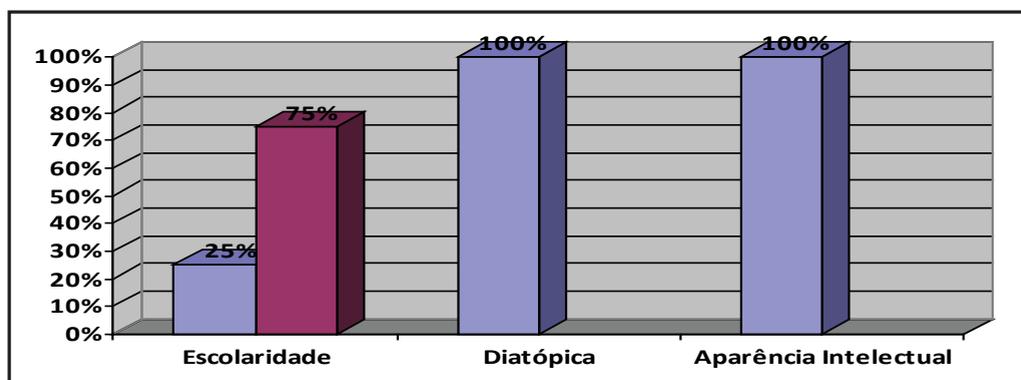
*Fragmento 4 - Se seus pais acham que há algo errado em namorar a esta altura do século, eles fazem o estilo supercaretas.[C82002P62]*

Assim, como ocorreu nos fragmentos 01 e 02, a partir da leitura dos dois fragmentos acima, foi solicitado aos mestrandos do Profletras que selecionassem expressões e/ou adjetivos de diferentes campos semânticos (diatópico, aparência física, escolaridade, aparência intelectual, profissional, estilo de vida/social, aparência/agradabilidade e avaliação social). Os itens selecionados foram:

**Gráfico 4:** sobre Fragmento 3<sup>14</sup>



**Gráfico 5:** sobre Fragmento 4<sup>15</sup>



<sup>14</sup> **Legenda:** Escolaridade: Ensino Fundamental Incompleto (35%), Ensino Médio Incompleto (65%) / Diatópico: Região Urbana da cidade (central) (40%) / Região Periférica da cidade (60%).

<sup>15</sup> **Legenda:** Escolaridade: Ensino Médio Incompleto (75%), Ensino Médio Completo (25%) / Diatópico: zona urbana (região central da cidade) (100%) / Aparência Intelectual: Inteligente (100%).

Como podemos observar nos gráficos 4 e 5, para o emprego de “ter” e “haver” com sentido existencial, alguns campos semânticos não foram selecionados pelos respondentes (são eles: aparência física, avaliação social, agradabilidade, profissional). Em relação aos resultados do uso do “ter” (fragmento 3), a maioria dos respondentes julgou o fragmento 3 como escrito por uma pessoa com o seguinte perfil: ter Ensino Médio Incompleto (65%), pertencer a zona urbana. Já em relação ao uso do “haver” (fragmento 4), predominantemente os respondentes avaliaram o fragmento 4 e atribuíram ao autor do texto o seguinte perfil: ter Ensino Médio Incompleto (75%), ser de zona central da cidade (100%) e ser “inteligente” (100%). Como podemos observar também, para o emprego de “haver”, os respondentes selecionaram mais itens e todos com julgamento positivo, como “ser inteligente”.

Acreditamos que esse resultado possa estar associado ao fato de, no caso desse fenômeno variável, o uso atrelado à prescrição normativa, reforçado pela escola (“haver” com sentido de existir), chamar mais atenção do que a variante “ter” com sentido de existir. Isso porque, se por um lado tal uso (do “ter”) não costuma ser prescrito pela gramática normativa ou, quando citado, acaba sendo associado a contextos de usos informais de fala e que devem ser evitados em contextos mais formais, sobretudo na escrita, por outro lado, vários estudos linguísticos já apontam a grande frequência e predominância do uso de ter com valor existencial no Português Brasileiro (CALLOU, AVELAR, 2001; DUTRA, 2000; VITORIO, 2010, OLIVEIRA; BARBOSA, 2014, entre outros). Ou seja, neste caso, diante de um fenômeno linguístico variável não estigmatizado e altamente produtivo na língua, a saliência recai para a variante padrão, colaborando para que os professores de língua portuguesa julgassem positivamente o uso de “haver” com sentido de existir, associando tal emprego ao perfil de “um bom aluno”.

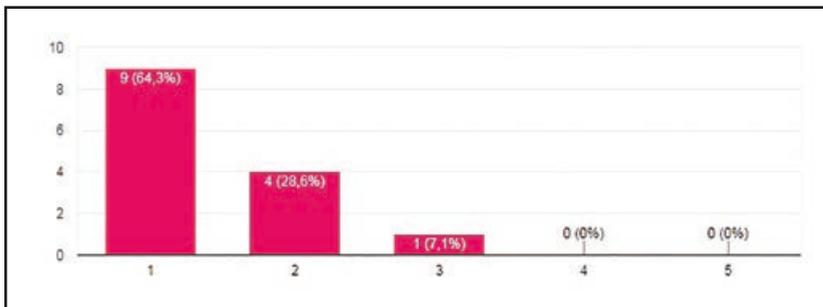
Entretanto, para os fragmentos 3 e 4, na questão “o autor é um bom aluno?”, todos os respondentes julgaram o produtor dos textos como sendo “bom aluno”, como aquele que “escreve um texto coerente” e três deles ressaltaram, para o fragmento 4, a importância do autor do texto ter utilizado o verbo “haver” e “saber usar a norma padrão”.

Por fim, temos os resultados principais da última parte do teste (parte III). Nessa etapa, os participantes da pesquisa foram solicitados a avaliar em quais contextos eles orientariam um determinado aluno “a escrever e/ou a utilizar as formas linguísticas” que apareciam em 04 fragmentos. Para isso, utilizamos, como já mencionado, uma escala de 1 a 5 (cf. CARDOSO, 2015), sendo

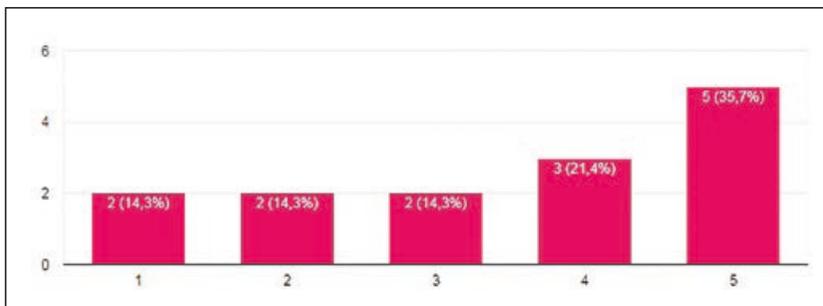
**1** contexto **menos monitorado** (ilustrado por uma situação de escrita menos formal) e **5** um contexto de **maior monitoramento** do falante (ilustrado por uma situação de escrita mais formal). Cabe mencionar que, nesses últimos itens do teste, dois respondentes não se manifestaram (alegaram desconfiguração da escala e dificuldade de preenchimento do formulário). Assim, nessa etapa tivemos apenas 14 respostas.

Para os fragmentos em que buscamos observar a avaliação dos mestrandos em relação à presença e à ausência de CV, tivemos os seguintes resultados:

**Gráfico 6:** fragmento com ausência de CV “3) Nós mudou porque a sala de cima era mais boa (M, 10, CB)



**Gráfico 7:** fragmento com presença de CV “4) Nós jogamos futebol, vôlei, basquete e tênis de mesa”. (M, 10, CB)

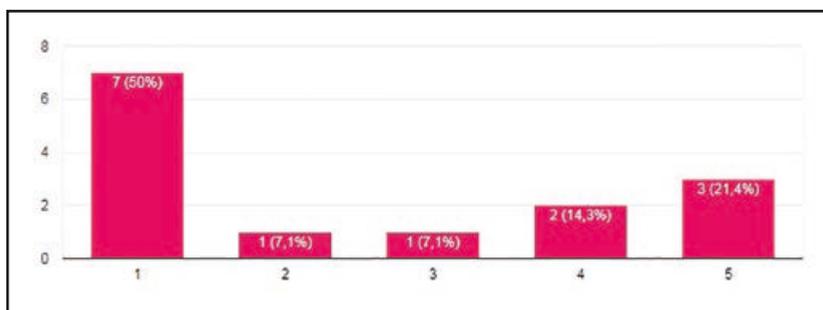


Ao compararmos os resultados dos gráficos 6 e 7, verificamos que os professores, ao avaliarem o fragmento com ausência de CV (gráfico 6), associaram-no ao contexto menos monitorado, que envolvia um grau de formalidade menor (escala 1 e 2, contexto de conversa escrita em *facebook* e *whatsapp* com amigo). Já em relação ao fragmento com presença de CV (gráfico 7), apesar dos professores terem apontado, em sua maioria, o contexto mais monitorado (escala 5 / 35,7%), alguns também associaram o fragmento e as estruturas da

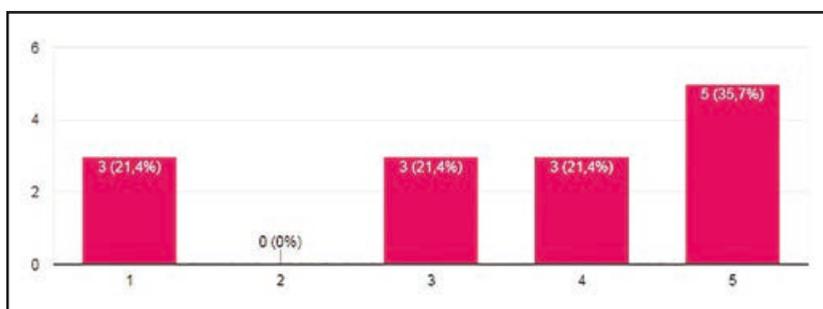
língua presentes neles, incluindo a concordância verbal, a um contexto menos monitorado (escalas 1 e 2, ambos com 14,3%). Tal resultado confirma a visão dos professores em relação à ausência da concordância verbal (associada a situações menos formais) e, conseqüentemente, podemos inferir que essa avaliação negativa do professor/mestrando do Profletras pode ser levada para sala de aula e transferida aos seus alunos da Educação Básica.

Já para a avaliação dos fragmentos com usos do “ter” e “haver” com sentido de existir, tivemos os seguintes resultados:

**Gráfico 8:** uso de TER “1) Tem pessoas que chegam e mudam os nossos planos”.



**Gráfico 9:** uso de HAVER “2) Há vários trabalhos para próxima semana”.



Como podemos observar pela análise dos gráficos 8 e 9, diferentemente do que ocorreu com a avaliação do fenômeno da CV (gráficos 6 e 7), o julgamento dos professores em relação aos fragmentos com o emprego de “ter” e “haver” com sentido de existir não se restringiu a uma das extremidades da escala (para menos ou para mais formal). Tivemos associações do uso de “ter/haver” tanto a contextos mais monitorados quanto para contextos menos monitorados. Sutilmente, 50% dos respondentes julgaram e associaram o emprego do verbo “ter”

a um contexto menos monitorado, ou seja, a uma situação de escrita menos formal. Esse resultado revela que, embora muitas pesquisas, como de Callou e Avelar (2001), já tenham apontado a preferência pelo uso de “ter” em diferentes amostras do Português Brasileiro – tanto na modalidade escrita quanto na falada –, ainda tal forma é associada a contexto de menor formalidade. Essa atitude negativa em relação ao uso de “ter” pode ser levada equivocadamente às aulas de língua portuguesa. Por outro lado, tivemos alguns respondentes que associaram o emprego de “ter” com sentido de existir (21,4%) a contextos mais monitorados, mais formais. Inversamente, o fragmento com presença do “haver” com sentido de existir foi associado por alguns professores a um contexto de maior formalidade (35,7%) e, simultaneamente, por outros professores, a um contexto de menor formalidade e menor monitoramento (21,4%).

Podemos inferir que essa oscilação na escala, em que os mestrandos do Profletras julgaram um mesmo fragmento com “haver” ou com “ter” tanto como menos formal como mais formal (respectivamente, escala de 1 e 5), se deve ao fato de ser um fenômeno menos suscetível à percepção da variação (menos saliente) e, conseqüentemente, recebe avaliações/julgamentos distintos.

## CONSIDERAÇÕES FINAIS

Neste capítulo, apresentamos os resultados de um teste piloto de atitude linguística aplicado a professores da Educação Básica da rede pública de ensino que são discentes ingressantes do ano de 2018 no Programa de Pós-Graduação Mestrado Profissional em Letras (Profletras) de duas unidades da região do Triângulo Mineiro.

Nosso teste, elaborado com fragmentos de produções escritas de alunos da Educação Básica, pautou-se, principalmente, em fenômenos variáveis do Português Brasileiro, sendo um supostamente mais estigmatizado (concordância verbal) e, outro, supostamente menos estigmatizado (os usos de “ter” e “haver” com sentido de existir).

Os resultados obtidos a partir da aplicação do teste nos deram pistas de como esses professores lidam com a variação linguística em sala de aula e como reagem frente a determinados fenômenos linguísticos variáveis, já que o julgamento sobre as variedades linguísticas que o falante emite, acaba refletindo, tal como mencionado anteriormente, na maneira como ele se comporta. Assim, nossa intenção também foi observar que tipo de atitude é mais comum entre esses professores de língua portuguesa: “uma atitude mais normativa, mais purista”

ou “uma atitude mais tolerante” (CARDOSO, 2015, p. 10). Nesse sentido, nossos resultados demonstraram que, em face a um fenômeno linguístico variável estigmatizado, como a concordância verbal, os professores tendem a ter atitudes linguísticas negativas quando a concordância não é marcada e positivas quando o uso da concordância é realizado de acordo com as prescrições da gramática normativa. Inclusive, alguns professores explicitaram a associação e a avaliação do “bom aluno” ao uso da concordância.

Entretanto, em face a um fenômeno linguístico variável não estigmatizado, como os usos de “ter e “haver” com sentido de existir, percebemos que as atitudes linguísticas dos professores se mostraram positivas diante dos fragmentos que representavam usos canônicos. Ou seja, tendo em vista que o uso de “ter” com sentido de existir é muito mais frequente no Português Brasileiro (uso não canônico), a saliência recai para a variante padrão que, ao ser identificada pelos professores, pode ter colaborado para que esses professores avaliassem mais positivamente o uso de “haver” com sentido de existir, associando tal uso com o perfil de “um bom aluno”.

Por fim, cabe destacar que nossa intenção com a elaboração, aplicação e análise de tal teste foi a de identificar e compreender as atitudes linguísticas de professores de língua portuguesa que estão ingressando em um curso de pós-graduação, cujo grande objetivo é contribuir para a formação continuada desses professores, de modo a capacitá-los para práticas de ensino mais sintonizadas com as contribuições da Linguística – em suas mais diversas subáreas – ao ensino de língua portuguesa.

Contribuições essas que já vêm sendo validadas por documentos oficiais atrelados ao ensino na Educação Básica no Brasil, como os Parâmetros Curriculares Nacionais (BRASIL, 1998) e, mais recentemente, a Base Nacional Curricular Comum (BNCC), cujas orientações para o Ensino Fundamental foram homologadas pelo Ministério da Educação em 2017 e, para o Ensino Médio, neste ano. Nesse sentido, acreditamos que nosso teste, após ajustes e aprimoramentos, pode se configurar como uma ferramenta avaliativa bastante profícua com vistas a subsidiar o planejamento de ações de conscientização linguística, formação continuada de professores e propostas de ensino do português. Isso porque, acreditamos que (re)conhecer as atitudes linguísticas dos professores de língua portuguesa que atuam na Educação Básica, na rede pública de ensino, pode subsidiar o planejamento de ações de conscientização acerca da heterogeneidade linguística do Português Brasileiro, além de contribuir para a elaboração de materiais didáticos de língua portuguesa mais sensíveis à variação linguística.

## REFERÊNCIAS

- AGHEYISI, R. N.; FISHMAN, J. A. Language attitude studies: A brief survey of methodological approaches. *Anthropological Linguistics* 12 (5), p. 137-157, 1970.
- AGUILERA, V. de A. Crenças e atitudes linguísticas: o que dizem os falantes das capitais brasileiras. In: *Estudos Linguísticos*, São Paulo, 37 (2): 105-112, maio-ago. 2008.
- ALMEIDA, J. L. de. Falar feio e falar bonito em *Quando as máquinas param*, de Plínio Marcos. In: PRETI, D. (org). *Cortesia verbal*. São Paulo: Editora Humanitas, 2008.
- BOTASSINI, J. A. M. A. Crenças e atitudes linguísticas quanto ao uso dos rótulos. *Signum*. Estudos de Linguagem, v. 12, p. 85-102, 2009.
- BOTASSINI, J. A. M. A. Importância dos estudos de crenças e atitudes para a Sociolinguística. *Signum*. v18. n1. 2015.
- CALLOU, D.; AVELAR, J. Sobre TER e HAVER em construções existenciais: variação e mudança no português do Brasil. *Gragoatá* 9, 2001, p. 85-114.
- CARDOSO, D. P. *Atitudes linguísticas e avaliações subjetivas de alguns dialetos brasileiros*. São Paulo: Blucher, 2015.
- CUBA, D. L.; BARBOSA, J. B. Crenças e atitudes linguísticas de alunos do ensino médio em escolas públicas de Uberaba. *Todas as Letras*, São Paulo, v. 17, n. 1, p. 73-90, jan./abr. 2015. Disponível em: < <http://editorarevistas.mackenzie.br/index.php/tl/article/view/7710/5299>>. Acesso em: 10 jul. 2019.
- CYRANKA, L. F. de M. Atitudes linguísticas de alunos de escolas públicas de Juiz de Fora – MG 2007. 174 f. Tese (Doutorado em Estudos Linguísticos). Instituto de Letras, Universidade Federal Fluminense, Niterói. 2007.

CYRANKA, L. F. de M. *Dos dialetos populares à variedade culta: a sociolinguística na escola*. Curitiba: Appris, 2011.

CYRANKA, F. de M. Avaliação das variantes: atitudes e crenças em sala de aula. In.: MARTINS, M. A.; VIEIRA, S. R.; TAVARES, M. A. (Org.) *Ensino de português e sociolinguística*. São Paulo: Contexto, 2014.

CYRANKA, L. F. de M; RONCARATI, C. Crenças de professores e alunos de português de escolas públicas de Juiz de Fora MG. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro II: contato linguístico, heterogeneidade e história*. Rio de Janeiro: FAPERJ/EDUFF, 2008.

DUTRA, C. Ter e haver na norma culta de Salvador. Dissertação (Mestrado em Letras) – Universidade Federal da Bahia, Bahia, 2000.

FREIRE, J. B. Variação, estilo, atitude e percepção linguística: o caso das laterais /ʎ/ e /l/ no falar paraibano. Tese de doutorado. 233f. João Pessoa, PB, 2016.

FREITAG, R. M. K. et al. A. Como os brasileiros acham que falam? Percepções sociolinguísticas de universitários do Sul e do Nordeste. *Todas as Letras*, n. 18, v.2, p. 64- -84, 2016.

FREITAG, R. M. K.; SANTOS, A. O. Percepção e atitudes linguísticas em relação às africadas pós-alveolares em Sergipe. In: SILVA, N. et al. *A fala nordestina : entre a sociolinguística e a dialetologia*. São Paulo : Blucher, 2016.

GARRETT, P. *Attitudes to language*. Cambridge: Cambridge University Press, 2010.

GARRETT, P.; COUPLAND, N.; WILLIAMS, A. *Investigating language attitudes*. University of Wales Press: Cardiff, 2003.

GÓMEZ MOLINA, J. R. Actitudes lingüísticas en una comunidad bilingüe y multidialectal: area metropolitana de Valencia. Anejo n. XXVIII dela Revista Cuadernos de Filología. Valencia, Universitat de Valencia, 1987.

GUESSI, R. R.; BARBOSA, J. B. Atitudes linguísticas de professores da rede pública de Uberaba-MG e o fenômeno variável de concordância verbal. Relatório de Iniciação Científica BIC-CNPQ. UFTM. 2017.

LABOV, W. *Principles of linguistic change*. Volume 2: Social factors. Oxford: Blackwell, 2001

LAMBERT, W. W., HODSON, R., GARDNER, R. & FILLENBAUM, S. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, vol. 60(1), 44–51, 1960.

LAMBERT, W. W.; LAMBERT, W. E. *Psicologia social*. Rio de Janeiro: Zahar, 1968.

MARINE, T. C.; BARBOSA, J. B. Crenças linguísticas de alunos do PROFLETRAS de universidades no Triângulo Mineiro. *Letrônica*. v. 10, n. 1. 2017.

MARINE, T. de C.; BARBOSA, J. B. Em busca de um ensino sociolinguístico de Língua Portuguesa no Brasil. *Signum: Estudos da Linguagem*, v. 19, n. 1, p. 185-215, jan. 2017. ISSN 2237-4876. Disponível em: <<http://www.uel.br/revistas/uel/index.php/signum/article/view/23161>>. Acesso em: 01 dez. 2019.

MORENO FERNÁNDEZ, F. *Principios de sociolingüística y sociología del lenguaje*. Barcelona: Ariel, 1998.

OLIVEIRA, A. A. S.; BARBOSA, J. B. TER e HAVER com sentido de existir em cartas da revista feminina *Capricho*. Relatório de Iniciação Científica BIC-CNPQ. UFTM. 2014.

SCHERRE, M. M. P. *Doa-se Lindos Filhotes de Poodle: Variação Linguística, Mídia e Preconceito*. São Paulo: Parábola Editorial, 2005.

SENE, M. G. Os desvios ortográficos em redações do Ensino Fundamental II: descrição, análise e atitudes linguísticas dos professores. Dissertação (Mestrado

em Linguística e Língua Portuguesa). Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras (Campus Araraquara). Araraquara, 2018.

SILVA, R. N. A. Variação ter/haver: contexto e ensino. In: HORA, D. D. et al. *Anais [do] Encontro Nacional de Ciências da Linguagem Aplicadas ao Ensino*. João Pessoa, UFPB, 2003. Disponível em: <[http://www.leffa.pro.br/tela4/Textos/Textos/Anais/ECLAE\\_II/varia%C3%A7%C3%A3o%20ter%20haver/principal.htm](http://www.leffa.pro.br/tela4/Textos/Textos/Anais/ECLAE_II/varia%C3%A7%C3%A3o%20ter%20haver/principal.htm)>. Acesso em jan. 2019.

TEODORO. M. A. Crenças linguísticas de professores de língua portuguesa da cidade de Uberaba. Dissertação (Mestrado em Educação). Universidade Federal do Triângulo Mineiro, Uberaba, 2018.

VITORIO, E. Ter/haver existenciais na escrita de alunos de 5ª e 6ª séries do ensino fundamental da cidade de Maracanaú/CE. 2006. Monografia (Especialização em Linguística e Ensino do Português) – Universidade Federal do Ceará, Ceará, 2006.



# LINGUISTIC ATTITUDES OF MASTER STUDENTS OF PROFLETRAS UNITS IN THE “TRIÂNGULO MINEIRO” REGION

*Juliana Bertucci Barbosa (UFTM-Uberaba)*  
*Talita de Cássia Marine (UFU-Uberlândia)*

Although the research of Labov (2008) – on the phonetic change in English spoken on the island of Martha’s Vineyard – already pointed out the importance of investigating the beliefs, attitudes and linguistic perception of speakers, in Brazil these works are still recent and on the rise such investigations (CYRANKA, 2007; CYRANKA; RONCARATI, 2008; AGUILERA, 2008; BOTASSINI, 2015, CUBA; BARBOSA, 2013; MARINE; BARBOSA, 2016, TEODORO, 2018, SENE, 2018, etc.). Studying the linguistic attitude of a group, for example, may help the researcher to understand the reactions of the speakers to the different varieties of a language, or their judgment in relation to different linguistic uses (GÓMEZ MOLINA, 1987, p.25). Moreover, through these investigations, we can understand how the linguistic attitude interferes in the process of constitution of the identity of a community (through its language). Thus, in this work, we seek to present results of a test of linguistic attitude applied to teachers of Basic Education in the public-school system. These teachers are entering the 2018 Program in the Professional Master’s Program in Letters (PROFLETRAS) of two units in the Triângulo Mineiro region, in Minas Gerais, Brazil. In order to construct the test, we are mainly based on Brazilian Portuguese variable phenomena, one being more stigmatized, such as the verbal

agreement phenomenon, and the other less stigmatized, like the verbs “ter” and “haver” with a sense “there is/are”. In general, our results showed that, faced with a stigmatized variable linguistic phenomenon, such as verbal agreement, teachers tend to have negative linguistic attitudes when agreement is not marked, and positive, when the use of agreement is performed according to prescriptions of normative grammar. In fact, some teachers made explicit the association and evaluation of the “good student” to the use of the agreement. However, in the face of a non-stigmatized variable linguistic phenomenon, such as the uses of “ter” and “haver” with a sense “there is/are”, we perceive that the linguistic attitudes of teachers were positive in the face of fragments representing canonical uses. Finally, it is important to highlight that our intention with the elaboration, application and analysis of such a test was to identify and understand the linguistic attitudes of Portuguese-speaking teachers who are entering a postgraduate course whose main objective is to contribute to the formation in order to train them for teaching practices more attuned to the contributions of Linguistics - in its most diverse subareas - to teaching Portuguese. In addition, we believe that understanding the linguistic attitudes of Portuguese-speaking teachers can support the planning of actions of linguistic awareness, ongoing teacher training and Portuguese teaching proposals.

## REFERENCES

- AGUILERA, V. de A. Crenças e atitudes linguísticas: o que dizem os falantes das capitais brasileiras. In: *Estudos Linguísticos*, São Paulo, 37 (2): 105-112, maio-ago. 2008.
- BOTASSINI, J. A. M. A. Importância dos estudos de crenças e atitudes para a Sociolinguística. *Signum*. v18. n1. 2015.
- CUBA, D. L.; BARBOSA, J. B. Crenças e atitudes linguísticas de alunos do ensino médio em escolas públicas de Uberaba. *Todas as Letras*, São Paulo, v. 17, n. 1, p. 73-90, jan./abr. 2015. Disponível em: < <http://editorarevistas.mackenzie.br/index.php/tl/article/view/7710/5299>>. Acesso em: 10 jul. 2019.
- CYRANKA, L. F. de M. Atitudes linguísticas de alunos de escolas públicas de Juiz de Fora – MG 2007. 174 f. Tese (Doutorado em Estudos Linguísticos). Instituto de Letras, Universidade Federal Fluminense, Niterói. 2007.

CYRANKA, L. F. de M.; RONCARATI, C. Crenças de professores e alunos de português de escolas públicas de Juiz de Fora MG. In: RONCARATI, C.; ABRAÇADO, J. (Org.). *Português brasileiro II: contato linguístico, heterogeneidade e história*. Rio de Janeiro: FAPERJ/EDUFF, 2008.

GÓMEZ MOLINA, J. R. Actitudes lingüísticas en una comunidad bilingüe y multidialectal: area metropolitana de Valencia. Anejo n. XXVIII dela Revista Cuadernos de Filología. Valencia, Universitat de Valencia, 1987.

SENE, M. G. Os desvios ortográficos em redações do Ensino Fundamental II: descrição, análise e atitudes linguísticas dos professores. Dissertação (Mestrado em Linguística e Língua Portuguesa). Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Ciências e Letras (Campus Araraquara). Araraquara, 2018.



---

# ENTREVISTA COM A SOCIOLINGUISTA DINAH MARIA ISENSEE CALLOU

*Marcos Luiz Wiedemer (UERJ)*

*Marcia dos Santos Machado Vieira (UFRJ)*

***O GT de Sociolinguística é um dos mais antigos no âmbito da ANPOLL. Quais aspectos destacaria ao traçar um panorama da atuação dele nesses anos nos cenários brasileiro e internacional? Que avaliação faz do percurso até então feito e do potencial de nossa contribuição nos cenários brasileiro e internacional?***

Diria que o GT de Sociolinguística, um dos mais antigos, engloba a Dialetologia, desde o início, e, mais tarde, a questão do contato (através da fusão com o GT de Bilinguismo) e da História da língua. Seu percurso está ligado à publicação do livro clássico de Labov sobre a linguagem de Nova York (1966) e à do texto de Weinreich *et alii* (1968) sobre a teoria da mudança linguística, que abriram novas perspectivas e constituem um marco no campo de estudo da variação e mudança linguística e na aproximação entre os temas referidos. A Sociolinguística despontou, no contexto dos estudos linguísticos brasileiros, como uma área desafiadora, dada a necessidade de compreender a realidade linguística de um país em que diferentes dimensões sociais se conjugam para a configuração de um quadro sociolinguístico complexo. O GT evoluiu rapidamente no Brasil e o desenvolvimento da Sociolinguística variacionista, com utilização de programas

computacionais de análise, introduzidos na UFRJ pelo professor Anthony Naro, foi fundamental para que se fizesse presente no cenário nacional e internacional, através da formação de Grupos de pesquisa pelo país afora e da divulgação dos resultados em Dissertações, Teses, artigos científicos e capítulos de livros.

***Considerando a proposição de uma agenda de trabalho, a médio e longo prazo, para o GT, quais destaques faria, principalmente, num cenário em que é necessário dar visibilidade a nossa pesquisa científica? Quais são as contribuições e as repercussões de nossa pesquisa no cenário internacional?***

Pensando em termos gerais, o GT cresceu muito nos seus mais de 30 anos de existência, sofrendo as reformulações necessárias à sua atuação. De início, o foco era observar a variação e mudança de forma mais sistemática e fazer um mapeamento sociolinguístico do Brasil. A médio prazo, fez-se necessário ampliar os objetivos e discutir as questões a partir de novos eixos, entre eles: princípios universais; interfaces teóricas; adequação de novos modelos para a área; política de expansão da LP na Europa, África e Américas; construção de novos *corpora* em diferentes regiões. Nossas pesquisas foram sendo ampliadas e ganharam maior visibilidade, tendo um papel fundamental no desenvolvimento da área no país. Em termos de agenda para o futuro, a proposta mais recente tem a ver com ‘avaliação’ e ‘atitudes’, a chamada ‘terceira onda’ que surgiu mais recentemente e passou a ter destaque nas discussões, no cenário nacional e internacional, já que implica o processo de construção de julgamentos subjetivos do falante em relação a sua própria língua e a do seu interlocutor, aí embutidos juízos de valor sobre as variedades linguísticas, que podem culminar em preconceito.

***Como o GT tem colaborado em prol da construção da Pós-Graduação no Brasil?***

Na medida em que corresponde a uma das linhas de pesquisa mais produtivas na maioria dos Programas de Pós-Graduação, com produção científica significativa, colabora efetivamente para o desenvolvimento da Pós-Graduação no país.

***Quais são as estratégias de inserção e visibilidade do GT de Sociolinguística na comunidade não-científica?***

Diria que de forma indireta, através da aceitação e respeito à diversidade linguística e pluralidade de normas – linguísticas e sociais – e às características sociais, culturais e históricas da região a que pertence, fundamental para a sua contextualização no letramento escolar. O Projeto de extensão, “*Ações de com-*

*bate ao preconceito linguístico*”, desenvolvido na UFRJ, pode ser uma estratégia para dar visibilidade aos estudos sociolinguísticos e à diversidade linguística e cultural.

***Como vê e como se configura o potencial de atuação interdisciplinar do nosso GT?***

O GT de Sociolinguística se configura como uma área que necessariamente reúne o linguístico e o social, ampliando sua ação interdisciplinar e suas interfaces teórico-metodológicas.

***Como vê a atuação do GT no que diz respeito à documentação de línguas e de culturas no Brasil e do Português dentro e fora do Brasil?***

O levantamento de *corpora* para o conhecimento da realidade linguística brasileira teve início no âmbito da Geolinguística e atingiu o ápice com o surgimento, a partir da década de 70 do século XX, de Projetos coletivos de pesquisa sociolinguísticos, como o *NURC*, *PEUL*, *Gramática do Português falado*, *Para a história do português brasileiro*, *ALIB*, *Concordância* e muitos outros, no decorrer dos anos, que contribuíram decisivamente para a expansão da área no cenário nacional e internacional.

***Que contribuição o GT pode oferecer ao processo de formação do leitor numa perspectiva de letramento científico e não científico?***

Além de contribuir para a descrição e explicação de fenômenos linguísticos, a sociolinguística também fornece subsídios para a área do ensino de línguas, sem negar a importância do conhecimento das teorias linguísticas que irão capturar padrões/regularidades e oferecer potencial explicativo para os fatos da língua. Com suas pesquisas baseadas na produção real dos indivíduos, a área da educação se enriquece, e, com a aplicação dos resultados das pesquisas sociolinguísticas na solução de problemas de letramento, surgem práticas mais efetivas de trabalho sobre a heterogeneidade linguística.

***Que palavra final destacaria para as novas gerações de pesquisa em Sociolinguística e Geolinguística?***

Apenas alguns lembretes:

Respeitem a diversidade! A variação existente hoje no português do Brasil é fruto da dinâmica populacional e da natureza do contato de diversos grupos étnicos e sociais;

- (i) Não desconheçam e/ou neguem a contribuição dos que nos precederam;
- (ii) Pesquisa e Ensino são indissociáveis se o indivíduo não quiser ser um mero repetidor. A busca do conhecimento não deve nem pode parar e cada avanço da ciência abre novos horizontes: “Não há assuntos esgotados, há homens esgotados diante de assuntos” (Ramón y Cajal);
- (iii) O trabalho em equipe é fundamental!

---

# OS ORGANIZADORES

## MARCIA DOS SANTOS MACHADO VIEIRA

Doutora (2001) e Mestre (1995) em Língua Portuguesa pela UFRJ e Bacharel e Licenciada em Português-Inglês pela UFRJ (1992), é Professora Associada IV do Departamento de Letras Vernáculas da UFRJ. Atua na Graduação e Pós-Graduação. Tem experiência em pesquisa sobre: variação e mudança de formas fonéticas (pretônicas) e morfossintáticas (formas de tratamento e de predicação); mudança por gramaticalização, lexicalização, mudança construcional ou construcionalização; predicação; impessoalização; auxiliaridade; temporalidade, aspectualidade e modalidade. Coordena, na UFRJ, o Projeto PREDICAR (Formação e expressão de predicados complexos), no qual se desenvolvem pesquisas em Linguística Funcional(-Cognitiva), Gramática das Construções e Sociolinguística. É membro do Grupo de Pesquisa “Discurso & Gramática (UFRJ). Desde 2016, integra a coordenação do eixo temático do GT de Sociolinguística da ANPOLL *Variação e Mudança Linguística*. Desde 2018 coordena, com Marcos Luiz Wiedemer, o GT de Sociolinguística da ANPOLL. Preside, também com esse colega, o Fórum Internacional em Sociolinguística.

## MARCOS LUIZ WIEDEMER

É doutor em Estudos Linguísticos pela UNESP, e curso ainda o doutorado sanduíche na Erfurt Universität (Alemanha), na área de Linguística, sob a orientação do Prof. Dr. Christian Lehmann, mestre em Linguística pela UFSC e licenciado em Letras (Português/Inglês) pela FURB. Atualmente é professor adjunto da Universidade do Estado do Rio de Janeiro (UERJ), atuando no curso de Letras (Português/Inglês) e no Programa de Pós-Graduação em Letras e Linguística (PPLIN), na Faculdade de Formação de Professores. É Coordenador

Geral do Programa de Pós-Graduação em Letras e Linguística (PPLIN), gestão 2018-2020. Membro pesquisador dos Grupos de Pesquisa “Discurso & Gramática (UFF) e “Estudos Sociofuncionalistas” (UFMS). Membro do Conselho da ANPOLL (Estudos Linguísticos). Desde 2018 coordena, com Marcia dos Santos Machado Vieira, o GT de Sociolinguística da ANPOLL. Preside, também com essa colega, o Fórum Internacional em Sociolinguística. Seu interesse de pesquisa envolve as seguintes áreas: Linguística Cognitivo-Funcional; Gramática das Construções e Sociolinguística.

---

# OS AUTORES

## ALBERT RILLIARD

Holder of a Doctorate in Cognitive Sciences from the Institut National Polytechnique de Grenoble (2000), of a Master in Language Sciences from the University Grenoble 3 (2000) and of a professional Thesis (“Habilitation à Diriger des Recherches”) in Computer Sciences from University Paris Sud (2014), Albert Rilliard is CNRS researcher at LIMSI (Orsay, France) and visiting professor at Rio de Janeiro Federal University (UFRJ, Brazil). He teaches phonetics and experimental data processing. His research interests are centered on variation in prosody, across languages and communication situations. Possibilities of communication mismatch across cultures, and the reasons rooting such problematic encounters are at the core of his questioning. He also works on the prosodic variation across Romance languages, applying objective metrics to help the measurement of observed differences, comparing with perception of spoken varieties – and the representation of dialects in different populations.

## ALESSANDRA DE PAULA

É Pesquisadora e Professora Adjunta de Língua Portuguesa e Filologia Românica no Departamento de Letras da Faculdade de Formação de Professores da Universidade do Estado do Rio de Janeiro (UERJ). Formou-se Bacharel e Licenciada em Português e Literaturas de Língua Portuguesa pela Universidade Federal do Rio de Janeiro (UFRJ) e é Doutora em Língua Portuguesa pela mesma instituição, com período de doutorado-sanduiche na Universidade de Lisboa e defesa da tese “Variação e mudança no vocalismo postônico medial

em português” (2015). Com apoio da FAPERJ, dedica-se ao estudo de aspectos fonológicos do português, com ênfase na fala fluminense, e investiga comunidades de fala do Leste Fluminense, no âmbito do projeto *Variação e mudança linguística em comunidades de fala do Leste Metropolitano do estado do Rio de Janeiro*. Tem publicações que abordam a variação e a mudança no vocalismo postônico não final do português – nas variedades brasileira e europeia – e os róticos – nas variedades moçambicana e santomense –, entre elas, “Vogais em contexto postônico não final em variedades do português: questões teóricas” (DE PAULA; BRANDÃO, 2015), “Panorama de aspectos sociolinguísticos do alteamento no vocalismo postônico medial em português” (DE PAULA, 2018) e “Róticos nas variedades urbanas santomense e moçambicana do Português” (BRANDÃO; DE PAULA, 2018).

## CATHERINE LEE

M.A., English, Texas Tech University, Lubbock, August 2012. Master’s Certificate, Linguistics, Texas Tech University, Lubbock, 2012. B.A., Geography, Texas Tech University, Lubbock, 2009.

## CHRISTINA ABREU GOMES

É professora titular do Departamento de Linguística e Filologia da Universidade Federal do Rio de Janeiro. Possui Doutorado em Linguística pela UFRJ e pós-doutorado na University of York. É bolsista de produtividade do CNPq. Orienta alunos de mestrado e doutorado do Programa de Pós-Graduação em Linguística da UFRJ. Sua atuação profissional tem focalizado temas desenvolvidos na Sociolinguística relativos à variação e mudança linguística, à aquisição e à percepção da variação socialmente indexada.

## CLÁUDIA REGINA BRESCANCINI

É Doutora em Letras pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS) e Mestre em Linguística pela Universidade Federal de Santa Catarina. Realizou estágio pós-doutoral no Language and Linguistic Department da University of York, Inglaterra. É professora adjunta do Programa de Pós-Graduação em Letras da PUCRS, onde desenvolve pesquisas na área de Sociolinguística, com ênfase em Sociofonética. É membro do Projeto Varsul (Variação Linguística na Região Sul do Brasil) desde 2005 e do Laboratório de Áudio e Fonética Acústica (Lafa) da Escola Politécnica da PUCRS desde

2006. Coordenou o GT de Sociolinguística da ANPOLL entre julho de 2014 e julho de 2018.

## CLINTON KAKELA AWAI

In University of Hawai'i at Mānoa, he has research that examines the relationship between sounds and social meanings.

## DANIELLE KELY GOMES

Professora Adjunta do Departamento de Letras Vernáculas da Universidade Federal do Rio de Janeiro, é Doutora e Mestre em Letras Vernáculas pela Universidade Federal do Rio de Janeiro (2012 e 2006, respectivamente), Bacharel e Licenciada em Português/Literaturas pela Universidade Federal do Rio de Janeiro (2003/2004). Foi professora da Educação Básica entre 2005 e 2013, atuando nas Secretarias Municipal e Estadual de Educação do RJ. Entre 2013 e 2014, foi professora Adjunta I do Departamento de Letras Clássicas e Vernáculas da Universidade Federal Fluminense. É uma das organizadoras de *Uma história de investigações sobre a língua portuguesa* (Blucher, 2018). Dedicar-se a pesquisas em teorias fonológicas, nas interfaces aquisição da linguagem/aprendizagem da escrita, contato linguístico e análises construtivas entre variedades do Português.

## DINAH MARIA ISENSEE CALLOU

Graduação em Letras Anglo-germânicas pela Universidade Federal da Bahia (1959), Mestrado em Língua Portuguesa pela Universidade de Brasília (1965), Doutorado em Letras Vernáculas - Área Língua Portuguesa) pela Universidade Federal do Rio de Janeiro (1980) e Pós-Doutorado em Linguística (Universidade da Califórnia/Santa Bárbara, 1994-1995). Pesquisador 1A (bolsa de produtividade em pesquisa) do CNPq e Professor Titular (1992) da Universidade Federal do Rio de Janeiro, onde atua desde 1966. Exerceu atividades na UFBA (1960-1962) e na UNB (1963-1965). Tem experiência na área de Fonética/Fonologia e Sintaxe, com ênfase em Sociolinguística e Linguística Histórica, com produção nos seguintes temas: variação e mudança, português do Brasil, fala culta carioca, sócio-história. Aposentada em 2008, continua atuando no Programa de Pós-Graduação em Letras Vernáculas, tendo recebido o título de Professor Emérito em 02/09/2010.

## **JONNY KIM (ALSO KNOWN BY THE NAME JUNG YUN KIM, IN KOREA)**

Linguist specializing in phonetics, phonology, psycholinguistics, and sociolinguistics. Ph.D. program in the Department of Linguistics at the University of Hawai'i at Mānoa in May, 2018. He works at the Hanyang Institute for Phonetics and Cognitive Sciences of Language (HIPCS) as a research assistant professor under the direction of Dr. Taehong Cho (PI).

## **JULIANA BERTUCCI BARBOSA**

Doutorado, fomentado pelo CNPq, em Linguística e Língua Portuguesa pela UNESP-Campus de Araraquara. Realizou, em Portugal, na Universidade de Lisboa, Faculdade de Letras (FLUL) e Centro de Linguística (CLUL), estágio de Doutorado PDEE fomentado pela CAPES. Tem experiência na área de Linguística, com ênfase em Sociolinguística e estudos descritivos no nível Sintático-Semântico. Também atua na constituição de banco de dados (montagem de *corpora*) e pesquisas variacionistas no Português Mineiro da cidade de Uberaba. Atualmente é professora do Departamento de Linguística e Língua Portuguesa da Universidade Federal do Triângulo Mineiro (UFTM-Uberaba) e líder do grupo de pesquisa GEVAR (Grupo de Estudos Variacionistas) e coordenadora e professora do PROFLETRAS. Desde julho de 2016 atua como membro permanente do GT de Sociolinguística da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística (ANPOLL). Suas pesquisas atuais estão concentradas na área de Sociolinguística com interface ao ensino de Língua Portuguesa e descrição do Português Brasileiro.

## **KATIE DRAGER**

Associate Professor in the Department of Linguistics at the University of Hawai'i at Mānoa specializing in Sociophonetics. She has been developing researches that examine the relationship between sounds and social meanings; she has been investigating, especially, how expectations about a speaker may influence how listeners interpret the sounds they produce. Her research relies on experimental methods to deal with phonetics, lexical access and social significance.

## **LIVIA OUSHIRO**

É professora do Instituto de Estudos da Linguagem/Unicamp, do Programa de Pós-Graduação em Linguística e coordenadora do Laboratório Variação,

Identidade, Estilo e Mudança (VARIEM). Fez sua graduação e pós-graduação na Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, onde se doutorou com uma tese a respeito do português paulistano. Tem interesse em estudos de produção e de percepções sociolinguísticas, sobretudo em grandes centros urbanos e em situações de contato dialetal. Em suas últimas pesquisas, tem analisado processos de mudança linguística na fala individual de migrantes nordestinos residentes nas cidades de São Paulo e de Campinas, levando em conta fenômenos de natureza fonética (p.ex., vogais pretônicas, realização de /r/ em coda silábica, pronúncia de /t, d/ antes de [i]) e morfossintática (concordância nominal e negação sentencial). Orienta trabalhos de Iniciação Científica, Mestrado e de Doutorado na área de Sociolinguística Variacionista sobre variação, identidade e difusão da mudança linguística. Também tem se preocupado com o desenvolvimento e a divulgação de novos métodos e questões na Sociolinguística Variacionista, principalmente pelo uso de novas ferramentas computacionais. Tem diversos trabalhos publicados em periódicos nacionais e internacionais e busca compartilhar tudo o que descobre em sua página na Internet (<https://oushiro.github.io/>).

## M. JOELLE KIRTLEY

PhD in Linguistics, University of Hawai'i at Mānoa, Dissertation- Language, Identity, and Non-Binary Gender in Hawai'i, Chair Katie Drager. M.A in Linguistics, University of Hawai'i at Mānoa, Thesis, Speech in the U.S. Military: A Sociophonetic Perception Approach to Identity and Meaning, Chair Katie Drager. B.A. in English, concentration in literature with a minor in music, 2006, John Brown University.

## MARCELO ALEXANDRE SILVA LOPES DE MELO

É professor Adjunto do Departamento de Linguística e Filologia da UFRJ. Possui Graduação em Direito e em Letras – Português Inglês, Mestrado e Doutorado em Linguística pela UFRJ. Tem experiência na área de Linguística, com ênfase em Sociolinguística, atuando principalmente na pesquisa dos seguintes temas: variação e mudança linguística, percepção da variação socialmente indexada e aquisição da escrita por população de baixa renda.

## MARCIA DOS SANTOS MACHADO VIEIRA

Doutora (2001) e Mestre (1995) em Língua Portuguesa pela UFRJ e Bacharel e Licenciada em Português-Inglês pela UFRJ (1992), é Professora Associada IV do Departamento de Letras Vernáculas da UFRJ. Atua na Graduação e Pós-Graduação. Tem experiência em pesquisa sobre: variação e mudança de formas fonéticas (pretônicas) e morfossintáticas (formas de tratamento e de predicação); mudança por gramaticalização, lexicalização, mudança construcional ou construcionalização; predicação; impessoalização; auxiliaridade; temporalidade, aspectualidade e modalidade. Coordena, na UFRJ, o Projeto PREDICAR (Formação e expressão de predicados complexos), no qual se desenvolvem pesquisas em Linguística Funcional(-Cognitiva), Gramática das Construções e Sociolinguística. É membro do Grupo de Pesquisa “Discurso & Gramática (UFRJ). Desde 2016, integra a coordenação do eixo temático do GT de Sociolinguística da ANPOLL *Variação e Mudança Linguística*. Desde 2018 coordena, com Marcos Luiz Wiedemer, o GT de Sociolinguística da ANPOLL. Preside, também com esse colega, o Fórum Internacional em Sociolinguística.

## MARCOS LUIZ WIEDEMER

É doutor em Estudos Linguísticos pela UNESP, e cursou ainda o doutorado sanduíche na Erfurt Universität (Alemanha), na área de Linguística, sob a orientação do Prof. Dr. Christian Lehmann. Mestre em Linguística pela UFSC e licenciado em Letras (Português/Inglês) pela FURB. Atualmente é professor adjunto da Universidade do Estado do Rio de Janeiro (UERJ), atuando no curso de Letras (Português/Inglês) e no Programa de Pós-Graduação em Letras e Linguística (PPLIN), na Faculdade de Formação de Professores. É Coordenador Geral do Programa de Pós-Graduação em Letras e Linguística (PPLIN), gestão 2018-2020. Membro pesquisador dos Grupos de Pesquisa “Discurso & Gramática (UFF) e “Estudos Sociofuncionalistas” (UFMS). Membro do Conselho da ANPOLL (Estudos Linguísticos). Desde 2018 coordena, com Marcia dos Santos Machado Vieira, o GT de Sociolinguística da ANPOLL. Preside, também com essa colega, o Fórum Internacional em Sociolinguística. Seu interesse de pesquisa envolve as seguintes áreas: Linguística Cognitivo-Funcional; Gramática das Construções e Sociolinguística.

## SILVIA FIGUEIREDO BRANDÃO

É Professora Titular de Língua Portuguesa (UFRJ), Doutora em Letras Vernáculas (1988, UFRJ) e atua desde 1978, na Área de Língua Portuguesa, no Departamento de Letras Vernáculas da Faculdade de Letras da UFRJ. Em 2009, realizou estágio pós-doutoral, com bolsa CAPES, no Centro de Linguística da Universidade de Lisboa. É bolsista de Produtividade em Pesquisa do CNPq desde 2008 e Cientista do Nosso Estado pela FAPERJ. No Programa de Pós-Graduação em Letras Vernáculas, desde 1990, desenvolve e orienta pesquisas no âmbito da Sociolinguística Variacionista e da Dialectologia, em especial nos campos da Fonética-Fonologia e da Morfossintaxe. Entre 1994 e 1996, coordenou o GT de Sociolinguística da ANPOLL, e, entre 2000 e 2004, o Projeto VARPORT (Análise Contrastiva de Variedades do Português), de Cooperação Internacional - CAPES-FCT (Brasil-Portugal). Em 2014, foi eleita Vogal da ALFAL, associação em que também coordena o Projeto 21. Autora de *A geografia linguística no Brasil* (São Paulo: Ática, 1991) e com publicações no Brasil e no exterior, organizou *Dois variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas* (2018) e, ainda, com a colaboração de Maria Antónia Mota, da Universidade de Lisboa, *Análise contrastiva de variedades do Português: primeiros estudos* (Rio de Janeiro: In-Fólio, 2003) e, em parceria com Silvia Rodrigues Vieira, *Morfossintaxe e ensino de Português: reflexões e propostas* (Fac. Letras-UFRJ, 2004) e *Ensino de gramática: descrição e uso* (São Paulo: Contexto, 2007).

## SILVIA RODRIGUES VIEIRA

Professora associada (nível III) do Departamento de Letras Vernáculas da Faculdade de Letras da UFRJ, Bolsista de produtividade 2 do CNPq (2015; 2018) e Pesquisadora FAPERJ, contemplada nos Programas Jovem Cientista do Nosso Estado 2011 (2012-2015) e Cientista do Nosso Estado 2014 (2015-2018). Possui Doutorado (2002) e Mestrado (1995) em Letras Vernáculas (Língua Portuguesa), além de Graduação em Português-Inglês pela Universidade Federal do Rio de Janeiro (1992). Atua nos Programas de Pós-Graduação em Letras Vernáculas e no Mestrado Profissional em Letras. Coordenadora de projetos de pesquisa nacional e de cooperação internacional (com destaque ao Grupo de trabalho da ALFAL - Projeto 21), organizadora de bancos de dados para análises linguísticas, sua produção bibliográfica conta com organização de livros e periódicos, capítulos publicados em livros, artigos em periódicos e trabalhos em anais de

congressos no âmbito nacional e internacional. Atua principalmente nas áreas da Sociolinguística, sobretudo no que se refere à descrição de variedades do Português, e do Ensino de gramática. No âmbito do GT de Sociolinguística da ANPOLL, coordena o Eixo relativo à Variação e Ensino, cujos trabalhos buscam (i) descrever a complexa rede de variedades do Português a serviço da sistematização e do estabelecimento de critérios para orientações relativas ao tema da norma-padrão e (ii) desenvolver metodologias para a abordagem pedagógica de regras variáveis.

## **TALITA DE CÁSSIA MARINE**

É professora e pesquisadora de Linguística e Língua Portuguesa do ILEEL (Instituto de Letras e Linguística da Universidade Federal de Uberlândia - UFU) e do Programa de Pós-Graduação Mestrado profissional em Letras (PROFLE-TREAS-UFU). Possui graduação em Letras (Licenciatura e Bacharelado) com habilitação em Português e Alemão pela UNESP/Araraquara. Desenvolveu pesquisa de Iniciação Científica fomentada pela Capes, como bolsista PET. Realizou Mestrado - fomentado pela Capes - e Doutorado - fomentado pelo CNPq - em Linguística e Língua Portuguesa na UNESP/Araraquara com estágio PDEE financiado pela Capes (“Doutorado-sanduíche”) na Faculdade de Letras da Universidade de Lisboa (FLUL) e no Centro de Linguística da mesma universidade (CLUL). É líder, na UFU, do Grupo de Estudos Variacionistas (GEVAR) - cadastrado na plataforma CNPq – e, desde julho de 2016, atua como membro permanente do GT de Sociolinguística da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística (ANPOLL). Suas pesquisas atuais estão concentradas na área de Sociolinguística com interface ao ensino de Língua Portuguesa.

## **VALÉRIA NETO DE OLIVEIRA MONARETTO**

Possui graduação em Letras pela Universidade Federal do Rio Grande do Sul (1988); Mestrado em Letras pela Universidade Federal do Rio Grande do Sul (1992); Doutorado em Linguística e Letras pela Pontifícia Universidade Católica do Rio Grande do Sul (1997); Pós-Doutorado na UFRJ (2013). Tem experiência na área de Linguística, com ênfase em Fonologia, Sociolinguística e Linguística Histórica, atuando principalmente nos seguintes temas, na graduação e na pós-graduação: variação e mudança fonológica, sob perspectiva sincrônica e diacrônica; teorias fonológicas modernas; pesquisas em linguística histórica. Foi

Vice-Coordenadora do GT de Sociolinguística da ANPOLL dos biênios no período 2014-2018. É professora Titular do Departamento de Linguística, Filologia e Teoria Literária da UFRGS desde 2015 e membro pesquisadora do Projeto Variação Linguística do Sul do País (VARISUL) desde 1995. Atualmente é a atual coordenadora deste Projeto na sede UFRGS, Regente da disciplina Linguística Histórica da graduação em Letras da UFRGS e Editora-Chefe do periódico “Cadernos do Instituto de Letras” do PPGLET/UFRGS, desde 2014.

