



DIADORIM

24
NÚMERO 1



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Diretora Adjunta de Pós-Graduação e Pesquisa

Profa. Dra. Aniela Improta França

Vice-diretora

Profa. Dra. Ana Regina Vaz Calindro

Coordenadora do Programa de Pós-Graduação em Letras Vernáculas

Profa. Dra. Eliete Figueira Batista da Silveira

Substituta Eventual do Coordenador

Profa. Dra. Violeta Virgínia Rodrigues

Comissão Deliberativa

Representantes Docentes

Língua Portuguesa

Profa. Dra. Célia Regina dos Santos Lopes

Profa. Dra. Marcia dos Santos Machado Vieira

Profa. Dra. Maria Eugenia Lammoglia Duarte (suplente)

Literatura Brasileira

Profa. Dra. Maria Lucia Guimarães de Faria

Profa. Dra. Anélia Montechiari Pietrani

Prof. Dr. Gilberto Araújo de Vasconcelos Júnior (suplente)

Literaturas Portuguesa e Africanas

Profa. Dra. Sofia Maria de Sousa Silva

Prof. Dr. Rodrigo Alexandre de Carvalho Xavier

Profa. Dra. Maria Teresa Salgado Guimarães da Silva (suplente)

Profa. Dra. Luciana dos Santos Salles (suplente)

Representantes Discentes

Daniel Veneri (Doutorando Representante de Literatura Brasileira)

Bruno Pereira dos Santos (Doutorando Representante de Literatura Brasileira (suplente))

Flávio Silva Corrêa de Mello (Doutorando Representante de Literaturas Portuguesa e Africanas)

Natércia Almeida Lacerda (Doutoranda Representante de Língua Portuguesa)

Secretário do Programa de Pós-Graduação em Letras Vernáculas

Renato Martins

Diretora da Faculdade de Letras

Profa. Dra. Sonia Cristina Reis

Vice-Diretor da Faculdade de Letras

Prof. Dr. Humberto Soares da Silva

Diretor Adjunto de Ensino de Graduação

Prof. Dr. Humberto Soares da Silva

Diretor Adjunto de Cultura e Extensão

Prof. Dr. Roberto de Freitas Junior

Diretor Adjunto de Administração e Finanças

Victor Hugo C. dos Santos

CENTRO DE LETRAS E ARTES (CLA)

Decania do Centro de Letras e Artes

Decana: Profa. Dra. Cristina Grafanassi Tranjan

Vice-decano: Prof. Dr. Osvaldo Luiz de Souza Silva

Reitora:

Profa. Dra. Denise Pires de Carvalho

Vice-reitor:

Prof. Dr. Carlos Frederico Leão Rocha

Sobre o primeiro número do volume 24 (2022)**Editora Chefe:**

Profa. Dra. Marcia dos Santos Machado Vieira, Universidade Federal do Rio de Janeiro, Brasil

Editoras Adjuntas:

Profa. Dra. Danielle Kely Gomes, Universidade Federal do Rio de Janeiro, Brasil

Profa. Dra. Eliete Figueira Batista da Silveira, Universidade Federal do Rio de Janeiro, Brasil

Editora Assistente de Literatura:

Profa. Dra. Sofia Maria de Sousa Silva, Universidade Federal do Rio de Janeiro, Brasil

Editoras Assistentes de Língua:

Profa. Dra. Ana Paula Quadros Gomes, Universidade Federal do Rio de Janeiro, Brasil

Profa. Dra. Leonor Werneck dos Santos, Universidade Federal do Rio de Janeiro, Brasil

Coordenação de Revisão:

Profa. Dra. Ana Paula Victoriano Belchor, Universidade Federal do Rio de Janeiro, Brasil

Equipe de Revisão:**Língua Portuguesa**

Vanessa Meireles, Université Paul-Valéry/Montpellier 3, França

Marcia dos Santos Machado Vieira, Universidade Federal do Rio de Janeiro, Brasil,

Anélia Montechiari Pietrani, Universidade Federal do Rio de Janeiro, Brasil

Língua Inglesa

Joan Simpson, Université Paul-Valéry - Montpellier 3, França

Marcia dos Santos Machado Vieira, Universidade Federal do Rio de Janeiro, Brasil

Maria Lucia Guimarães de Faria, Universidade Federal do Rio de Janeiro, Brasil

Equipe Técnica:

Rafael Laplace de Andrade, Agoodigital, Rio de Janeiro, Brasil

Miguel R. Amorim Neto, Universidade Federal do Rio de Janeiro, Brasil

Design e Diagramação:

Rafael Laplace de Andrade, Agoobook, Rio de Janeiro, Brasil

Rodrigo Pereira Martins, Agoobook, Rio de Janeiro, Brasil

Redação:

Programa de Pós Graduação em Letras Vernáculas - Faculdade de Letras da UFRJ

Av. Horácio de Macedo, 2151 - sala F321, Ilha do Fundão, Cidade Universitária

CEP 21941-917 - Rio de Janeiro/RJ

E-mail: posvernaculas@letras.ufrj.br; revdiadorim@letras.ufrj.br

<https://revistas.ufrj.br/index.php/diadorim>

Diadorim: Revista de Estudos Linguísticos e Literários - Volume 24, número 1 de 2022.
Rio de Janeiro: UFRJ, Programa de Pós-Graduação em Letras Vernáculas.

Organizadores da edição (v.24, n.1 de 2022):

Anélia Montechiari Pietrani, Universidade Federal do Rio de Janeiro, Brasil

Godofredo de Oliveira Neto, Universidade Federal do Rio de Janeiro, Brasil

Marcia dos Santos Machado Vieira, Universidade Federal do Rio de Janeiro, Brasil

Rita Olivieri-Godet, Université Rennes 2, França

Sophie Sarrazin, Université Paul-Valéry/Montpellier 3, França

Vanessa Meireles, Université Paul-Valéry/Montpellier 3, França

ISSN: 1980-2552.

E-ISSN: 2675-1216.



SUMÁRIO



Volume 24, número 1 de 2022

Prefácio / Preface

Prefácio

- 12** *Marcia dos Santos Machado Vieira, Danielle Kely Gomes e Eliete Figueira Batista da Silveira*

Preface

- 14** *Marcia dos Santos Machado Vieira, Danielle Kely Gomes and Eliete Figueira Batista da Silveira*

Apresentação/Presentation - Dossiê de Língua/Language Dossier

- 16** **Variação em línguas românicas: Occitano e Português**
Vanessa Meireles, Marcia dos Santos Machado Vieira e Sophie Sarrazin

- 24** **Variation in Romance Languages: Occitan and Portuguese**
Vanessa Meireles, Marcia dos Santos Machado Vieira and Sophie Sarrazin

Entrevista(s) / Interview(s)

- 32** **Análise de dados linguísticos – Entrevista com Adriana Picoral**
Adriana Picoral, Marcia dos Santos Machado Vieira, Vanessa Meireles, Ravena Beatriz de Sousa Teixeira e Mariana Gonçalves da Costa

- 40** **Place sociale et variation diasystemique de L'Occitan – Interview avec Hervé Lieutard**
Hervé Lieutard, Marcia dos Santos Machado Vieira e Vanessa Meireles

- 53** Lugar social e variação diassistêmica do Occitano – Entrevista com
Hervé Lieutard
Hervé Lieutard, Marcia dos Santos Machado Vieira e Vanessa Meireles

Dossiê de Língua / Language Dossier

- 66** Aquisição/ensino do Português Europeu por/a falantes arabófonos em
marrocos: especificidades e elementos de resposta
Maria Antónia Mota, Habiba Naciri
-

- 83** Variação linguística, ensino de Língua Portuguesa e atitudes linguísticas
Juliana Bertucci Barbosa, Daiana Lombardi de Cuba
-

- 112** Subject expression in spoken varieties of Portuguese
Aline Maria Bazenga
-

- 143** Pronominal subject in Alagoas: a case of change in progress
Maria Auxiliadora da Silva Cavalcante
-

- 161** A variação nós e a gente em Fortaleza na segunda década dos anos 2000:
fatores linguísticos
*Maylle Lima Freitas, Lorena da Silva Rodrigues, Hugo Leonardo Gomes
dos Santos*

180 O efeito do paralelismo linguístico sobre a disputa entre o indicativo e o subjuntivo na formação do imperativo de 2ª pessoa do singular no Português Mineiro histórico (séculos XIX e XX)
Luiz Fernando de Carvalho

203 A concordância nominal como um marcador estilístico de construção da persona estereotipada do acadêmico da UFSC
Ariete Helena Holz Nunes, Sabrina Vieira Teixeira

233 Valor diminutivo do sufixo -ão em Português
Graça Rio-Torto

251 Variação no campo semântico “pesos e medidas”: um estudo do léxico Catarinense
Antonio Luiz Gubert, Vanderci de Andrade Aguilera

276 Botar o filho na escola: os verbos *botar* e *colocar* no falar de Fortaleza-CE.
Camilo Murilo Alves de Lavor, Aluiza Alves de Araújo e Rakel Beserra de Macedo Viana

298 O tipo de sintagma na presença de artigo definido antes de possessivos
Manoel Siqueira

315 Vogais médias pretônicas no Português do Libolo (Angola): comparações com o Português Brasileiro
Paloma Moreira Freire, Flaviane Romani Fernandes Svartman

336 Aspectos da entoação na fala de pacientes com esquizofrenia
Waldemar Ferreira Netto, Marcus Vinícius Moreira Martins e Ana Cristina Aparecida Jorge

Resenha(s) de Língua

366 **EJA e idosos: caminhos de inclusão**
Maria Cecília de Magalhães Mollica

Apresentação/Presentation - Dossiê de Literatura/Literature Dossier

370 **Tecendo textos, abrindo trilhas: escritoras indígenas contemporâneas**
*Anélia Montechiari Pietrani, Godofredo de Oliveira Neto e
Rita Olivieri-Godet*

375 **Weaving texts, opening trails: contemporary indigenous women writers**
*Anélia Montechiari Pietrani, Godofredo de Oliveira Neto and
Rita Olivieri-Godet*

Entrevista(s)

379 **Autoria feminina no cordel indígena: entrevista com Auritha Tabajara**
Janda Montenegro

382 **Respirar e resistir: entrevista com Graça Graúna**
*Marta Passos Pinheiro, Guilherme Trielli Ribeiro e Viviane de Cássia Maia
Trindade*

388 Autoexpressão e resistência em *Coração na aldeia, pés no Mundo*, de
Auritha Tabajara
Ana Maria de Carvalho

408 Memória e ancestralidade em poemas de *Metade cara, metade máscara*,
de Eliane Potiguara
Débora Francisca de Lima

423 *Heart Berries: a Memoir* (2018), de Terese Marie Mailhot, e a cura de
palavras maltratadas
Fernanda Vieira de Sant'Anna

450 A importância da mulher indígena para a ancestralidade, identidade e
tradição nos poemas “O segredo das mulheres” e “Mulher!”, de Eliane
Potiguara
Maria do Carmo Moreira de Carvalho, Algemira de Macêdo Mendes

466 Gotas de histórias: Márcia Kambeba registra a memória ancestral
dos Omágua/Kambeba
Paulo Marcelino dos Santos, Elizabeth Gonzaga de Lima

487 Mulher-pássaro e vento ancestral: narrativas indígenas de Graça
Graúna e Mariela Tulián
Randra Kevelyn Barbosa Barros

505 *Flor da mata: reverberações estéticas de conhecimentos indígenas da*
natureza em haicais de Graça Graúna
Rinah de Araújo Souto, José Hélder Pinheiro Alves

518 Tessituras do tempo: resenha de *Fios do Tempo: (quase haikais)*, de Graça Graúna
José André Souza Silva e Rosivânia dos Santos

523 *Weiyamî – mulheres que fazem sol: o ritual poético de Sony Ferseck*
Olivieri-Godet



PREFÁCIO

A comissão editorial da *Diadorim: revista de estudos linguísticos e literários* (periódico de acesso aberto) tem a honra de apresentar mais um volume, composto por dois dossiês: língua e literatura.

O primeiro número do volume 24 da *Diadorim: revista de estudos linguísticos e literários* é publicado pelo Programa de Pós-Graduação em Letras Vernáculas da Faculdade de Letras da Universidade Federal do Rio de Janeiro. Nos dois dossiês, congrega duas temáticas definidas pelos pesquisadores docentes que foram convidados a atuarem como editores do número, bem como contribuições, na forma de entrevistas, resenhas e artigos, oriundos de olhares e pesquisas presentes em diferentes instituições no Brasil e no exterior.

Seguindo as diretrizes editoriais recomendadas, os editores do número assumiram a função de, recebidos os artigos, submetê-los ao processo de avaliação duplo cego entre pesquisadores das áreas temáticas. Dessa prática, resultou este volume que reúne, ao todo: 4 entrevistas com pesquisadores atuantes nas áreas relacionadas neste número, 20 artigos/ensaios, além dos textos de apresentação dos dois dossiês e 3 resenhas.

O dossiê temático de estudos linguísticos do primeiro número do volume 24 da *Diadorim: revista de estudos linguísticos e literários* reúne visões sobre as variações e variedades do português e também da língua occitana e revela como são frutíferas e diversificadas as experiências na ciência linguística. Aborda questões relacionadas a espaços de formação de pesquisadores e professores, espaços de ensino e aprendizagem (considerando as línguas materna e não materna), bem como espaços de comunicação científica e impacto dos resultados da pesquisa na sociedade. Esse dossiê é organizado por **Vanessa Meireles (Universidade Paul-Valéry Montpellier 3)**, **Marcia dos Santos Machado Vieira (Universidade Federal do Rio de Janeiro)** e **Sophie Sarrazin (Universidade Paul-Valéry Montpellier 3)**.

O dossiê temático de estudos literários do primeiro número do volume 24 da *Diadorim: revista de estudos linguísticos e literários* reúne textos sobre literatura, sobre o contexto atual do movimento, observado nas Américas, de afirmação das culturas dos povos originários, sobre as obras escritas por mulheres ameríndias em suas tensas relações com as culturas nacionais institucionalizadas e sobre questões relacionadas a essa produção, atravessada pela subjetividade

feminina e pelo desejo para a emancipação literária e social. Esse dossiê é organizado por **Anélia Montechiari Pietrani (Universidade Federal do Rio de Janeiro)**, **Godofredo de Oliveira Neto (Universidade Federal do Rio de Janeiro, Academia Brasileira de Letras)** e **Rita Olivieri-Godet (Universidade de Rennes 2 / Instituto Universitário da França)**.

Agradecemos, imensamente, aos editores deste número e, especialmente, aos pesquisadores pareceristas que aceitaram o convite daqueles para avaliarem os artigos submetidos aos dois dossiês, pela colaboração, a um só tempo, valorosa, dado o esforço inestimável de lidar com mais uma tarefa acadêmica em meio a tantas a administrar, e essencial ao diálogo, que se prega na academia e que pode ser tão proveitoso ao trabalho dos autores e editores.

Temos, então, mais um número em que olhares de língua e literatura se somam e nos proporcionam uma rica experiência de leitura. Torcemos para que o leitor tire proveito do que aqui é reunido ou a partir daqui possa ser potencializado.

Boa leitura! E bom proveito deste número em futuros espaços de interlocução!

Editoras da Diadorim

Marcia dos Santos Machado Vieira (editora-chefe)

Danielle Kely Gomes (editora adjunta)

Eliete Figueira Batista da Silveira (editora adjunta)





PREFACE

The editorial committee of *Diadorim: revista de estudos linguísticos e literários* (an open access journal) is honored to present another volume, comprising two dossiers: language and literature.

The first issue of volume 24 of *Diadorim: revista de estudos linguísticos e literários* is published by the Postgraduate Program in Vernacular Letters of the Faculty of Letters of the Federal University of Rio de Janeiro. In both dossiers, it brings together two themes defined by the research professors who were invited to act as editors of the issue, as well as contributions, in the form of interviews, reviews and articles, from the perspectives and researches present in different institutions in Brazil and abroad.

Following the recommended editorial guidelines, the issue's editors assumed the role of, after receiving the articles, submitting them to the double-blind evaluation process among researchers in the thematic areas. This practice resulted in this volume that brings together, in total: 3 interviews with researchers working in the areas listed in this issue, 20 articles/essays, in addition to the presentation texts of the two dossiers and 3 reviews.

The thematic dossier on linguistic studies in the first issue of volume 24 of *Diadorim: revista de estudos linguísticos e literários* brings together views on variations and varieties of Portuguese and also on the Occitan language and reveals how fruitful and diverse experiences in linguistic science are. It deals with issues related to spaces for training researchers and teachers, spaces for teaching and learning (taking into consideration both mother and non-mother languages), as well as spaces for scientific communication and the impact of research results on society. This dossier is organized by **Vanessa Meireles (Paul-Valéry University Montpellier 3)**, **Marcia dos Santos Machado Vieira (Federal University of Rio de Janeiro)** and **Sophie Sarrazin (Paul-Valéry University Montpellier 3)**.

The thematic dossier on literary studies in the first issue of volume 24 of *Diadorim: revista de estudos linguísticos e literários* brings together texts on literature, on the current context of the movement, observed in the Americas, to affirm cultures of original peoples, on works written by Amerindian women in their tense relations with institutionalized national cultures and on issues related to this production, crossed by the female subjectivity and the desire

Diadorim, Rio de Janeiro, vol. 24, número 1, p. 14 - 15, 2022.



for literary and social emancipation. This dossier is organized by **Anélia Montechiari Pietrani (Federal University of Rio de Janeiro), Godofredo de Oliveira Neto (Federal University of Rio de Janeiro, Brazilian Academy of Letters) and Rita Olivieri-Godet (Université Rennes 2 / Institut Universitaire de France).**

We are immensely grateful to the editors of this issue and, especially, to the peer researchers who accepted their invitation to evaluate the articles submitted to the two dossiers, for their valuable collaboration, given the invaluable effort of dealing with yet another academic task amidst so many to manage, and essential to the dialogue, which is defended in academia and which can be so beneficial to the work of authors and editors.

We have, then, another issue in which language and literature perspectives add up and provide us with a rich reading experience. We hope that the reader takes advantage of what is gathered here or from here that it can be leveraged.

Good reading! And take advantage of this issue in future spaces for dialogue!

Diadorim Editors

Marcia dos Santos Machado Vieira (editor-in-chief)

Danielle Kely Gomes (associate editor)

Eliete Figueira Batista da Silveira (associate editor)





**VARIAÇÃO EM LÍNGUAS ROMÂNICAS:
OCCITANO E PORTUGUÊS**

**VARIATION IN ROMANCE LANGUAGES:
OCCITAN AND PORTUGUESE**

Vanessa Meireles¹

Marcia dos Santos Machado Vieira²

Sophie Sarrazin³

Este dossiê temático, publicado no primeiro número do volume 24 da *Diadorim: revista de estudos linguísticos e literários*, do Programa de Pós-Graduação em Letras Vernáculas da Universidade Federal do Rio de Janeiro, reúne olhares sobre variações e variedades do português e também olhar sobre a língua occitana. Articula-se a ações que vêm sendo construídas nesse programa em prol de ampla internacionalização da ciência linguística promovida a partir do espaço acadêmico brasileiro.

Uma delas é a gestada a partir do projeto franco-brasileiro *VariaR – Variação em línguas românicas* (<https://variari.wixsite.com/variari>), cuja equipe de pesquisadores, atuantes em diferentes universidades, é coordenada por docentes vinculadas a programas de pesquisa e pós-graduação com sede na Universidade Paul-Valéry Montpellier 3/unidade *Recherches Sur les Suds et les Orients* e na Universidade Federal do Rio de Janeiro/Programa de Pós-Graduação em Letras Vernáculas. O dossiê contém, ao todo, 19 textos, entre treze artigos submetidos à avaliação duplo cega e aprovados por pareceristas especialistas do Brasil, Portugal, Suécia, Estados Unidos e França, assim como duas entrevistas com dois pesquisadores, uma com pesquisadora do continente americano/Estados Unidos e outra, em francês e em português, com pesquisador do continente europeu/França, e uma resenha de livro brasileiro feita por pesquisadora brasileira.

1 Professora do Grupo ReSO da Universidade Paul Valéry Montpellier 3, coordenadora do Projeto VariaR – Variação em Línguas Românicas. E-mail: vanessa.meireles@univ-montp3.fr.

2 Professora Associada do Departamento de Letras Vernáculas e do Programa de Pós-Graduação em Letras Vernáculas da UFRJ, coordenadora do Projeto VariaR – Variação em Línguas Românicas. E-mail: marcia@letras.ufrj.br.

3 Professora do Grupo ReSO da Universidade Paul Valéry Montpellier 3. E-mail: sophie.sarrazin@univ-montp3.fr.

As pesquisas expostas nos artigos aqui reunidos focalizam a língua portuguesa, que é o centro da atenção do Projeto Capes PrInt do Programa de Pós-Graduação em Letras Vernáculas intitulado *Vozes e escritas nos diferentes espaços da língua portuguesa*. Esta é outra frente de ação em prol da internacionalização da ciência linguística brasileira.

O occitano, que normalmente tem sido minorizado nas descrições, é outra língua românica objeto de nossa atenção e, por conseguinte, de observações feitas numa das entrevistas concedidas às organizadoras do dossiê.

Em linhas gerais, os textos deste dossiê revelam como são frutíferas e diversas as experiências da ciência linguística. Há artigos sobre vários objetos linguísticos, mais precisamente sobre objetos das áreas de fonética-fonologia, prosódia, léxico, morfologia, sintaxe e texto/discurso. Há textos a tematizarem questões relativas a espaços de formação de pesquisadores e professores, espaços de ensino-aprendizagem, perspectivando língua materna ou não materna, assim como espaços de comunicação científica e impacto dos resultados de investigações na sociedade. Lê-se, nesse sentido, desde entrevista sobre o processo metodológico de recorte e análise de dados linguísticos, haja vista o interesse em rumos de ciência aberta, até resenha que dá uma mostra das implicações sociais de pesquisas em Letras e Linguística com atenção especial a um processo de alfabetização inclusivo e também a um dos grupos de sujeitos em que essa reverbera, o de idosos, tendo em vista rumos de educação cidadã. Os autores dos textos são oriundos de diferentes lugares: Estados Unidos, França, Brasil (regiões nordeste, sudeste, sul), Portugal (Lisboa, Coimbra e Ilha da Madeira) e Marrocos. As ênfases na abordagem de variação em que se centra cada texto também são múltiplas: Sociolinguística Variacionista sincrônica, Sociolinguística Histórica, Sociolinguística Aplicada, Dialectologia Pluridimensional ou Geossociolinguística. Há igualmente estudos de variação que se alinham aos trabalhos feitos nas chamadas primeira e terceira ondas de práticas científicas na ampla área da Teoria de Variação e Mudança.

O dossiê deste número de periódico brasileiro foi concebido para acolher a expressão científica gerada a partir de observações empíricas sobre variação em línguas românicas no Brasil e no mundo. Não obstante, os artigos voltam-se à língua portuguesa. Por um lado, os artigos recebidos, tratando apenas de usos da língua portuguesa, confirmam a vitalidade de vozes e escritos sobre ela na comunidade científica que trabalha dentro do Brasil ou noutros países. Por outro, em alguma medida também sinalizam que é importante investir, em território brasileiro, em ações e interações de descrição sobre línguas românicas e sobre comparações entre português e outra(s) língua(s) românica(s), perfil de investimento científico a nortear a equipe do Projeto VariaR. Emergem, por conseguinte, campos de atuação a explorar e agenda de trabalho a construir, para que possamos reunir vozes e escritos com perfil comparativo que evidenciem divergências, mas, principalmente, convergências relativas a um conhecimento

multilinguístico que, se supõe, envolve graus de diassistematicidade. Isso pode ensejar articulação de esforços que têm sido feitos e podem ser encontrados, a título de ilustração, em publicações centradas (i) numa ou mais variedades da língua portuguesa (MARTINS; ABRAÇADO, 2015, BRANDÃO, 2018, DE PAULA et al, 2018, para citar alguns trabalhos), (ii) no *continuum* afro-brasileiro, como o da obra *The Portuguese language continuum in Africa and Brazil* (LÓPEZ; GONÇALVES; AVELAR, 2018), e (iii) em línguas românicas como, por exemplo, *The Oxford guide to the romance languages* (LEDGEWAY; MAIDEN, 2016), o *Manual of Romance Sociolinguistics* (AYRES-BENNETT; CARRUTHERS, 2018) e o livro *Variação em português e em outras Línguas Românicas* (MACHADO VIEIRA; MEIRELES, 2022), este organizado em torno do estudo de variações em português principalmente e incluindo outras línguas românicas (espanhol, francês e italiano).

A perspectiva que norteou este dossiê está em sinergia, portanto, com o esforço coletivo de continuar a fazer não somente o mapeamento do português em suas diversas manifestações nacionais/dialetais contemporâneas, mas também de situá-lo no conjunto de línguas românicas, a fim de estabelecer um terreno comum de generalizações e especificidades, bem como reverberar em ações de formação de recursos humanos a atuarem em ciência e educação. É, portanto, com esse espírito de fomentar uma rede de colaborações para o desenho da variação em línguas românicas mais próximo possível da realidade linguística pluricêntrica que se passa a apresentar cada contribuição deste dossiê. Os primeiros textos são os das entrevistas, que antecedem os artigos e, por fim, a resenha.

Na entrevista concedida pela Professora Doutora Adriana Picoral, que atua na Universidade do Arizona, estão em foco temas relativos à ciência da análise de dados linguísticos, ciência da informação e ciência computacional. Atualmente investigando variação linguística no Português do Brasil na rede social Twitter e variação no Inglês por aprendizes, a pesquisadora trata de métodos de análise linguística relativos ao enfoque variacionista da Sociolinguística em particular e da Linguística de Corpus. Destaca a importância dos métodos estatísticos para a generalização de análises linguísticas confiáveis e, sobretudo, a importância de esses métodos serem reproduzíveis em novas investigações, uma garantia de qualidade da análise em si e um dos princípios *FAIR* em sintonia com demandas de ciência aberta. Outro destaque trazido pela professora é a necessidade de se fomentar a formação de linguistas na área das ciências exatas. Ela também fala do seu papel na iniciativa global *Women in data Science*, para a inclusão de mulheres em espaços de pouca representatividade como é o caso do setor de informática.

A entrevista concedida em francês por Hervé Lieutard, Professor Titular de Linguística Occitana na Universidade Paul-Valéry – Montpellier 3, foi traduzida para o português, para uma maior acessibilidade ao público não francófono. Nessa entrevista, o professor especialista

em análise da variação dialetal do occitano, língua falada atualmente em três Estados diferentes (França, Espanha e Itália), fornece um breve panorama do estatuto e das dificuldades da pesquisa em occitano a partir do lugar social dessa língua ao longo de sua história em território francês. O professor destaca o declínio social do uso dessa língua ao longo do tempo, fruto de uma política linguística de longa data que minorizou essa língua na sociedade francesa, ao ponto de ser considerada como uma língua regional na França, sem estatuto oficial, apesar de seu desenvolvimento histórico, uso, contato e variação linguística importantes. O linguista também discorre sobre os processos que têm contribuído para a revitalizar, como o crescente interesse da comunidade acadêmica e a ampliação da oferta de ensino.

Aquisição/Ensino do português europeu por/a falantes arabófonos em Marrocos: especificidades e elementos de resposta é o primeiro artigo. Foi escrito pelas investigadoras Maria Antónia Mota (da Universidade de Lisboa, Portugal) e Habiba Naciri (da Universidade Mohammed V-Rabat, Marrocos). Centra-se no enquadramento, dos pontos de vista acadêmico e sociolinguístico, da primeira Licenciatura em Estudos Portugueses no mundo árabe, cujo início ocorre no ano letivo de 2009-2010. As autoras destacam a cooperação institucional que enseja a proposição de tal Licenciatura, o multilinguismo social a caracterizar a população estudantil de Marrocos interessada em Português, as condições em que Português constitui uma língua adicional, a construção de materiais didáticos para a aquisição tardia por arabófonos de Português, na sua variedade europeia, como língua estrangeira mediante o desenvolvimento do Projeto MAPEAR. E ilustram, com fatos linguísticos, alguns aspectos detectados nas produções escritas e orais dos estudantes, bem como empreendem uma reflexão sobre o que dados linguísticos sugerem quanto a interlínguas e ações em espaços de ensino do Português.

O artigo *Variação linguística, ensino de língua portuguesa e atitudes linguísticas*, de Juliana Bertucci Barbosa (professora da Universidade Federal do Triângulo Mineiro, colaboradora do Programa de Pós-Graduação em Linguística e Língua Portuguesa UNESP/Araraquara) e Daiana Lombardi de Cuba (pesquisadora da Universidade Estadual Paulista/Araraquara, professora da Educação Básica), trata da questão do trabalho com a variação linguística em espaços de ensino. Tendo em mente os Parâmetros Curriculares Nacionais (PCN), as autoras discorrem sobre os novos caminhos da Sociolinguística no campo do ensino e também apresentam resultados sobre um teste de atitudes linguísticas aplicado a alunos da Educação Básica.

Aline Maria Bazenga (da Universidade da Madeira) é autora do artigo *A expressão do sujeito em variedades do Português*. Nele, ao tratar de dados de um fenômeno morfossintático, a pesquisadora lida com a relação entre variedades faladas do Português e diversidade de perfis histórico-sociais e culturais das comunidades que usam essas variedades. Ela centra sua atenção na expressão variável do sujeito (nulo e pronominal), examinando-a com base em contributos oriundos de modelos teóricos sobre a variação linguística no âmbito da Gramática Generativa, Tipologia Linguística e Sociolinguística.

O sujeito pronominal em Alagoas: um caso de mudança em progresso, escrito por Maria Auxiliadora da Silva Cavalcante (da Universidade Federal de Alagoas), tematiza a alegada preferência por realizações plenas em detrimento dos sujeitos nulos que consta dos estudos de amostras de fala realizados na região sudeste. Valendo-se de uma amostra de dados de Alagoas relativos ao espaço de sujeitos pronominais examinada com base em orientações da Teoria de Variação Linguística, a autora reúne evidências de um processo de mudança no sentido de uma inclinação maior à realização dos sujeitos pronominais.

Em *A variação nós e a gente em fortaleza na segunda década dos anos 2000: fatores linguísticos*, de Maylle Lima Freitas, Lorena da Silva Rodrigues e Hugo Leonardo Gomes dos Santos (todos da Universidade Federal do Ceará), os autores focam em resultados do estudo da variação dos pronomes sujeito *nós* e *a gente* numa comunidade de fala do Português Brasileiro: falantes de Fortaleza-CE com ensino superior na segunda década dos anos 2000. A investigação descrita nesse artigo foi realizada com base nos pressupostos da Teoria de Variação e Mudança e a partir de entrevistas do Projeto Descrição do Português Oral Culto de Fortaleza (PORCUFORT).

Luiz Fernando de Carvalho (da Universidade Federal de Minas Gerais), em *O efeito do paralelismo linguístico sobre a disputa entre o indicativo e o subjuntivo na formação do imperativo de 2ª pessoa do singular no português mineiro histórico (séculos XIX e XX)*, reúne evidências empíricas, com base em amostras históricas (cartas mineiras oitocentistas e novecentistas autógrafas), para mostrar que a variação do imperativo é sensível aos paralelismos sintático e fônico e, então, tematizar a questão dos condicionamentos linguísticos na dinâmica de variação.

Em *Concordância nominal como um marcador estilístico de construção da persona estereotipada do acadêmico da UFSC*, Ariele Helena Holz Nunes e Sabrina Vieira Teixeira (ambas da Universidade Federal de Santa Catarina), à luz de orientações sociolinguísticas, especialmente as vinculadas à chamada terceira onda, e de registros viabilizados por entrevistas nos moldes labovianos variacionistas da primeira onda, mostram a importância da concordância nominal como marcador estilístico da identidade da persona acadêmica.

O artigo *Valor diminutivo do sufixo -ão em português*, de Graça Rio-Torto (da Universidade de Coimbra), traz uma descrição do uso deste sufixo -ão em português tendo em conta sua dualidade semântica de ‘diminuição’ e ‘aumentação’. A partir de fontes do português de ocorrência do sufixo com valor diminutivo, ainda que menos comum, também presente em outras línguas românicas, a autora analisa as condições estruturais e históricas da ocorrência deste padrão românico na língua portuguesa, apoiando-se no arcabouço teórico da morfologia construcional e cognitivista.

Em *Variação no campo semântico “pesos e medidas”*: um estudo do léxico catarinense, Antonio Luiz Gubert (do Instituto Federal de Santa Catarina) e de Vanderci de Andrade Aguilera (da Universidade Estadual de Londrina) mostram resultados da investigação sobre o léxico dos “pesos e medidas” usado na região de Xanxerê (Santa Catarina) e em municípios dos entornos. Para tanto, valem-se de orientações teóricas da Dialetoologia Pluridimensional ou Geossociolinguística e exploram variação diacrônica, diageracional, diastrática e diassexual.

Botar o filho na escola: os verbos botar e colocar no falar de Fortaleza-CE é outro artigo que põe em evidência uma variação do nível lexical. Foi escrito por Camilo Murilo Alves de Lavor, Aluiza Alves de Araújo e Rakel Beserra de Macedo Viana (todos da Universidade Estadual do Ceará). Os pesquisadores analisam sociolinguisticamente o efeito de grupos de fatores linguísticos e extralinguísticos sobre a alternância entre os verbos em usos oriundos do banco de dados NORPOFOR (Norma Oral do Português de Fortaleza).

No artigo *O tipo de sintagma na presença de artigo definido antes de possessivos*, Manoel Siqueira (da Universidade Federal de Sergipe) explora empiricamente, com base na observação sociolinguística de dados oriundos de amostras de fala de estudantes universitários da Universidade Federal do Sergipe, a inclinação detectada em outras investigações brasileiras quanto à relação entre a variável tipo de sintagma e as variantes presença e ausência de artigo.

Vogais médias pretônicas no português do Libolo (Angola): comparações com o português brasileiro é o artigo escrito por Paloma Moreira Freire e Flaviane Romani Fernandes Svartman, pesquisadoras da Universidade de São Paulo. As autoras objetivam contribuir para a descrição da variedade africana do português do Libolo. Então, apresentam, a partir da observação de alguns aspectos, como se realizam as vogais pretônicas do português do Libolo, Angola (PL), comparando o que detectam aos resultados obtidos com as descrições encontradas na literatura sobre as vogais pretônicas do português brasileiro (PB). Para essa comparação, exploram dados oriundos de gravações de fala espontânea produzidas por seis falantes angolanos do município do Libolo. As autoras destacam estes resultados: no PL, diferentemente do PB, as vogais médias-baixas e o fenômeno de *abaixamento vocálico* em posição pretônica são ausentes; como no português brasileiro, as vogais médias-altas pretônicas do português do Libolo podem sofrer o fenômeno de *alçamento vocálico*; e as variedades paulista e gaúcha do PB mostram-se, até o momento, as mais semelhantes ao PL.

Em *Aspectos da entoação na fala de pacientes com esquizofrenia*, de autoria de Waldemar Ferreira Netto (Universidade de São Paulo), Marcus Vinícius Moreira Martins (Universidade do Estado de Minas Gerais) e Ana Cristina Aparecida Jorge (Universidade de São Paulo), uma metodologia para caracterização da entoação na fala de sujeitos com esquizofrenia é desenvolvida, o que pode facilitar o diagnóstico dos pacientes. Para tanto, os autores realizaram

um experimento com 10 pacientes com esquizofrenia e 10 sujeitos-controle, utilizando o aplicativo ExProsodia e cálculo da entropia de Shannon para as frases analisadas. Os resultados obtidos indicam que a entoação da fala de sujeitos com esquizofrenia apresenta forte tendência à diminuição na quantidade de informação que pode ser detectada por meio da variação de F0.

Em *EJA e idosos: caminhos de inclusão*, Maria Cecília de Magalhães Mollica (da Universidade Federal do Rio de Janeiro) oferece-nos uma resenha do livro *Alfabetização e idosos – um tema urgente no Brasil*, de autoria de Kátia Abreu (da Universidade do Estado do Rio de Janeiro) e Maria Carlota Rosa (da Universidade Federal do Rio de Janeiro). Ao sintetizar aspectos de seus dezoito capítulos ligados a sistemas de escrita, aprendizagem de leitura, alfabetização via Educação de Jovens e Adultos, a pesquisadora convida à leitura da obra a comunidade acadêmico-profissional da área de Letras e Linguística interessada em movimentos e percursos da Educação no Brasil e na articulação entre o mundo teórico-científico e políticas públicas.

Este dossiê é o resultado da reunião dessas ricas contribuições que propiciam ao leitor traçar um valioso panorama em Linguística calcado em usos variáveis das línguas portuguesa e occitana e num amplo repertório de olhares teóricos, metodológicos, analíticos e descritivos do fenômeno de variação sob observação em diferentes espaços acadêmicos no Brasil e noutros países dos continentes americano e europeu.

As organizadoras deste dossiê são muito gratas à equipe editorial do periódico pela oportunidade e à equipe de diagramação do número. São muito agradecidas ao conselho consultivo e aos pareceristas que atuaram, com dedicação e tempo, no exame de cada artigo submetido, para que pudessem chegar à seleção dos que aqui estão. Também são gratas aos docentes, discentes e demais profissionais do Programa de Pós-Graduação em Letras Vernáculas pelo apoio na ampla divulgação deste dossiê, do belo número que o acolhe. Agradecem, ainda, à equipe da Universidade Paul-Valéry Montpellier 3, pelo frutífero diálogo, desde a concepção do Projeto VariaR. Convidam, por fim, o leitor a apreciar o dossiê, na certeza de que há aqui ricos caminhos de reflexão, descrição e/ou ação a reverberarem noutros rumos e espaços de ciência e educação. Desejam-lhe, assim, bom proveito.

Referências

AYRES-BENNETT, Wendy; CARRUTHERS, Janice (ed.). *Manual of Romance Sociolinguistics*. Berlin: De Gruyter, 2018, xi + 793 pp.

BRANDÃO, Sílvia F. (ed.). *Duas variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher Open Access, 2018.

DE PAULA, Alessandra; GOMES, Danielle K.; SILVEIRA, Eliete F. B.; MACHADO VIEIRA, Marcia dos S; VIEIRA, Sílvia R. (eds.). *Uma História de Investigações sobre a Língua Portuguesa: Homenagem a Sílvia Brandão*. São Paulo: Blucher Open Access. 2018. <https://www.blucher.com.br/livro/detalhes/uma-historia-de-investigacoes-sobre-a-lingua-portuguesa-homenagem-a-silvia-brandao-1406>

LEDGEWAY, Adam; MAIDEN, Martin (ed.). *The Oxford guide to the romance languages*. Oxford; New York: Oxford University Press, 2016.

LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de (ed.). *The Portuguese language continuum in Africa and Brazil*. Amsterdam: John Benjamins Publishing Company, 2018.

MACHADO VIEIRA, Marcia dos Santos; MEIRELES, Vanessa (ed.). *Variação em português e em outras Línguas Românicas*. São Paulo: Blucher Open Access, 2022. <https://www.blucher.com.br/variacao-em-portugues-e-em-outras-linguas-romanicas>

MARTINS, Marco Antonio; ABRAÇADO, Jussara. (ed.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015.



**VARIATION IN ROMANCE LANGUAGES:
OCCITAN AND PORTUGUESE**

**VARIAÇÃO EM LÍNGUAS ROMÂNICAS:
OCCITANO E PORTUGUÊS**

Vanessa Meireles¹

Marcia dos Santos Machado Vieira²

Sophie Sarrazin³

This thematic dossier, published in the first issue of volume 24 of *Diadorim: revista de estudos linguísticos e literários* – a journal compiled by the Postgraduate Program in Vernacular Letters at the Federal University of Rio de Janeiro –, brings together views on variations and varieties of Portuguese and also on the Occitan language. It is linked to initiatives that have been taken within this program in favor of the broad internationalization of linguistic science promoted in the Brazilian academic sphere.

One of these initiatives stems from the Franco-Brazilian project *Variar* – *Variation in Romance Languages* (<https://variav.wixsite.com/variav>), whose team of researchers, working at various universities, is led by professors linked to research and post-graduate programs based at Paul-Valéry University Montpellier III/*Recherches Sur les Suds et les Orient*s unit and at the Federal University of Rio de Janeiro/Postgraduate Program in Vernacular Letters. The dossier contains a total of 19 texts, including 13 articles submitted for double-blind evaluation and approved by expert reviewers from Brazil, Portugal, Sweden, the United States and France. It also contains two interviews with two researchers, one with a researcher from the American continent/the United States and another, written in French and Portuguese, with a researcher from the European continent/France, as well as a review of a Brazilian book by a Brazilian researcher.

1 Professor part of the ReSO Group, at Paul Valéry University Montpellier 3, coordinator of the *Variar* Project – Variation in Romance Languages. E-mail: vanessa.meireles@univ-montp3.fr.

2 Associate Professor in the Department of Vernacular Letters and the Postgraduate Program in Vernacular Letters at UFRJ, coordinator of the *Variar* Project – Variation in Romance Languages. E-mail: marcia@letras.ufrj.br.

3 Professor part of the ReSO Group, at Paul Valéry University Montpellier 3. E-mail: sophie.sarrazin@univ-montp3.fr.

The research presented in the articles gathered here focuses on the Portuguese language, which is the subject of the Capes PrInt Project run by the Postgraduate Program in Vernacular Letters, entitled *Voices and Writings in the Different Spaces of the Portuguese Language*. This is another initiative taken in favor of the internationalization of Brazilian linguistic science.

Occitan, language whose importance has often been minimized in descriptions, is another Romance language which forms part of our study interests and, consequently, is the subject of observations made in one of the interviews granted to the organizers of the dossier.

Generally speaking, the texts in this dossier reveal how fruitful and diverse experiences in linguistic science are. They include articles on various linguistic concepts, more precisely concepts in the areas of phonetics/phonology, prosody, lexicon, morphology, syntax, and text/discourse. The dossier also includes texts on issues related to spaces for training researchers and teachers, spaces for teaching and learning (taking into consideration both mother and non-mother languages), as well as spaces for scientific communication and the impact of research results on society. In this way, a wide variety of texts can be read, from an interview about the methodological process of clipping and analyzing linguistic data, given the interest in open science movements, to a review that shows the social implications of research in Letters and Linguistics. The latter pays special attention to an inclusive literacy education process, and also to one of the groups most affected, the elderly, with a view to citizen education. The authors of the texts come from various places: the United States, France, Brazil (Northeast, Southeast, South), Portugal (Lisbon, Coimbra and Madeira Island) and Morocco. The texts are also centered around multiple approaches to variation: Synchronic Variationist Sociolinguistics, Historical Sociolinguistics, Applied Sociolinguistics, Pluridimensional Dialectology and Geosociolinguistics. There are also studies of variation that align with the work done in the so-called first and third waves of scientific practices in the broad area of the Theory of Variation and Change.

The dossier in this issue of this Brazilian journal was conceived to accommodate the scientific expression generated from empirical observations on variation in Romance languages in Brazil and in the world. However, the articles focus on the Portuguese language. On the one hand, the articles received, dealing only with uses of the Portuguese language, confirm the vitality of what has been said and written about it in the scientific community that works within Brazil or in other countries. On the other hand, to some extent they also indicate that it is important to invest, on Brazilian territory, in actions and interactions relating to the description of Romance languages and comparisons between Portuguese and one or several other Romance language(s). This constitutes a scientific investment guideline for the VariaR Project team. Thereby, fields of action to be explored and a work agenda to be built have emerged, allowing us to gather speeches and writings with a comparative profile that show divergences, but, mainly, convergences related to a multilingual knowledge that is assumed to involve degrees of diasystemacity. This can create a link with previous efforts which can be found, by way

of illustration, in publications focused on (i) one or more varieties of the Portuguese language (MARTINS; ABRAÇADO, 2015, BRANDÃO, 2018, DE PAULA et al, 2018, to name a few works), (ii) the Afro-Brazilian *continuum*, such as *The Portuguese Language Continuum in Africa and Brazil* (LÓPEZ; GONÇALVES; AVELAR, 2018), and (iii) Romance languages, for example *The Oxford Guide to the Romance Languages* (LEDGEWAY; MAIDEN, 2016), the *Manual of Romance Sociolinguistics* (AYRES-BENNETT; CARRUTHERS, 2018) and the book *Variation in Portuguese and Other Romance Languages* (MACHADO VIEIRA; MEIRELES, 2022). The latter is organized mainly around the study of variations in Portuguese, but it includes other Romance languages (Spanish, French and Italian).

The vision that guided this dossier is in synergy, therefore, with the collective effort to continue not only to map Portuguese in its various contemporary national/dialectal manifestations, but also to place it in the Romance language family, in order to establish a common ground of generalizations and specificities, as well as to influence the training of human resources for work in science and education. Each contribution to this dossier is therefore presented in the spirit of fostering a network of collaborations for the mapping of variation in Romance languages as closely as possible to a pluricentric linguistic reality. The first texts are the interviews, which are followed by the articles and, finally, the review.

The interview given by Professor Adriana Picoral, who works at the University of Arizona, focuses on topics related to the science of linguistic data analysis, information science and computational science. Currently investigating linguistic variation in Brazilian Portuguese on the social network Twitter and variation in English spoken by learners, the researcher deals in particular with methods of linguistic analysis related to the variationist approach of Sociolinguistics, as well as to Corpus Linguistics. She highlights the importance of statistical methods for the generalization of reliable linguistic analyses and, above all, the importance of these methods being reproducible in new investigations, a guarantee of quality of the analysis itself and one of the *FAIR* principles in line with the requirements of open science. Another aspect highlighted by the professor is the need to encourage the training of linguists in exact sciences. She also talks about her role in the global initiative *Women in Data Science*, for the inclusion of women in spaces where they are under-represented, such as the information technology sector.

The interview given in French by Hervé Lieutard, Titular Professor of Occitan Linguistics at Paul-Valéry University Montpellier III has been translated into Portuguese for greater accessibility to the non-French-speaking public. In this interview, the professor, who specializes in the analysis of the dialectal variation of Occitan, a language currently spoken in three different countries (France, Spain and Italy), provides a brief overview of the status and difficulties of research into Occitan, based on the social position of this language throughout its history on French territory. The professor highlights the social decline of the use of this language over

time, the result of a long-standing linguistic policy that has reduced its importance in French society, to the point of its being considered a regional language with no official status, despite the importance of its historical development, usage, linguistic contacts and linguistic variation. The linguist also talks about the processes that have contributed to revitalizing the language, such as the growing interest of the academic community and the expansion of the teaching offer.

Acquisition/Teaching of European Portuguese by/to Arabic Speakers in Morocco: Specificities and Response Elements is the first article. It was written by the researchers Maria Antónia Mota (from the University of Lisbon, Portugal) and Habiba Naciri (from Mohammed V-Rabat University, Morocco). It focuses on the framework, from an academic and sociolinguistic point of view, of the first University Degree in Portuguese Studies in the Arab world, which began in the academic year of 2009-2010. The authors highlight the institutional cooperation that gave rise to the proposition of such a degree, the social multilingualism that characterizes the student population of Morocco interested in Portuguese, the conditions in which Portuguese constitutes an additional language, and the construction of teaching materials for late acquisition of European Portuguese as a foreign language by Arabic speakers through the development of the MAPEAR Project. They also illustrate, using linguistic facts, some aspects detected in the written and oral productions of students, as well as reflecting on what linguistic data suggests regarding interlanguages and actions in Portuguese teaching spaces.

The paper *Linguistic Variation, Portuguese Language Teaching and Linguistic Attitudes*, by Juliana Bertucci Barbosa (professor at the Federal University of Triângulo Mineiro, contributor to the UNESP/Araraquara Postgraduate Program in Linguistics and Portuguese Language) and Daiana Lombardi from Cuba (researcher at São Paulo State University/Araraquara, teacher of Basic Education), deals with the issue of working with linguistic variation in teaching spaces. Bearing in mind the National Curriculum Parameters (PCN), the authors discuss the new paths of Sociolinguistics in the field of education and also present results of a test of linguistic attitudes applied to Basic Education students.

Aline Maria Bazenga (from the University of Madeira) is the author of the article *Subject Expression in Spoken Varieties of Portuguese*. In this article, when describing data linked to a morphosyntactic phenomenon, the researcher deals with the relationship between spoken varieties of Portuguese and the diverse historical, social and cultural profiles of the communities that use them. She focuses on the variable expression of the (null and pronominal) subject, examining it based on contributions from theoretical models of linguistic variation within the scope of Generative Grammar, Linguistic Typology and Sociolinguistics.

The Pronominal Subject in Alagoas: A Case of Change in Progress, written by Maria Auxiliadora da Silva Cavalcante (from the Federal University of Alagoas), focuses on the alleged preference for full realizations to the detriment of null subjects that appears in the studies of speech samples carried out in the Southeast region. Using a sample of data from Alagoas

concerning the space of pronominal subjects examined based on guidelines from the Theory of Linguistic Variation, the author gathers evidence of a process of change in the direction of a greater inclination towards the realization of pronominal subjects.

In *Variation of the Subjective Pronouns nós and a gente in the Second Decade of the 2000s: Linguistic Factors*, written by Maylle Lima Freitas, Lorena da Silva Rodrigues and Hugo Leonardo Gomes dos Santos (all from the Federal University of Ceará), the authors focus on the results of the study of the variation of subject pronouns *nós* and *a gente* in a Brazilian Portuguese speech community: speakers from Fortaleza-CE with an university education in the second decade of the 2000s. The investigation described in this article was carried out based on the assumptions of the Theory of Variation and Change and on interviews from the Description of Cultivated Oral Portuguese in Fortaleza Project/Projeto Português Oral Culto de Fortaleza (PORCUFORT).

Luiz Fernando de Carvalho (from the Federal University of Minas Gerais), in *The Effect of Linguistic Parallelism on the Dispute between the Indicative and the Subjunctive in the Formation of the 2nd Person Singular Imperative in Historical Portuguese from Minas Gerais (19th and 20th Centuries)*, gathers empirical evidence, based on historical samples (19th century and 20th century original handwritten letters from Minas Gerais), to show that imperative variation is sensitive to syntactic and phonic parallelisms. He then discusses the issue of linguistic conditioning in the dynamics of variation.

In *Nominal Agreement as a Marker Style of Construction of the Stereotyped Persona of the UFSC Academic*, Ariele Helena Holz Nunes and Sabrina Vieira Teixeira (both from the Federal University of Santa Catarina), show the importance of nominal agreement as a stylistic marker of the identity of the academic persona in light of sociolinguistic orientations, especially those linked to the so-called third wave, and of records made possible by interviews using the Labovian variationist models of the first wave.

The paper *Diminutive Value of the Suffix -ão in Portuguese*, written by Graça Rio-Torto (from the University of Coimbra), brings a description of the use of the suffix *-ão* in Portuguese, taking into account its semantic duality as a diminutive and an augmentative. Using sources of Portuguese showing occurrences of the suffix with the diminutive value which, although less common, is also present in other Romance languages, the author analyzes the structural and historical conditions of the occurrence of this Romance pattern in the Portuguese language, based on the theoretical framework of constructional and cognitivist morphology.

In *Variation in the Semantic Field of “Weights and Measures”*: *A Study of the Lexicon of Santa Catarina*, written by Antonio Luiz Gubert (from the Federal Institute of Santa Catarina) and Vanderci de Andrade Aguilera (State University of Londrina), the authors show the results of research on the lexicon of “weights and measures” used in the region of Xanxerê (Santa Catarina) and in surrounding municipalities. To do so, they make use of theoretical orientations

from Pluridimensional Dialectology and Geosociolinguistics, and explore diachronic, diagenational, diastratic and diasexual variation.

Botar *the Son to School: The Verbs botar and colocar in the Speech of Fortaleza-CE* is another article that highlights a variation on the lexical level. It was written by Camilo Murilo Alves de Lavor, Aluiza Alves de Araújo and Rakel Beserra de Macedo Viana (all from the State University of Ceará). The researchers carry out a sociolinguistic analysis of the effect of groups of linguistic and extralinguistic factors on the alternation between verbs in uses derived from the NORPOFOR (Norma Oral do Português de Fortaleza / Oral Norm of Portuguese from Fortaleza).

In the article *The Type of Phrase in the Presence of the Definite Article Before Possessives*, Manoel Siqueira (from the Federal University of Sergipe) empirically explores, based on the sociolinguistic observation of data from speech samples of university students at the Federal University of Sergipe, the bias detected in other Brazilian investigations regarding the relationship between variable syntagm type and variants categorized by presence and absence of the article.

Pretonic mid-vowels in Libolo Portuguese (Angola): comparisons with Brazilian Portuguese is a paper written by Paloma Moreira Freire and Flaviane Romani Fernandes Svartman (researchers from the University of São Paulo). The authors aim to contribute to the description of the African variety of Portuguese in Libolo. Then, based on the observation of certain aspects, they present how the pretonic vowels of Portuguese from Libolo/Angola (PL) are pronounced, comparing what they detect to the results obtained from descriptions found in literature on pretonic vowels in Brazilian Portuguese (BP). For this comparison, they explore data from spontaneous speech recordings produced by six Angolan speakers from the municipality of Libolo. The authors highlight the following results: in PL, unlike in BP, mid-low vowels and the phenomenon of vowel lowering in the pretonic position are absent. As in Brazilian Portuguese, the pretonic mid-high vowels of Portuguese from Libolo can be subject to the phenomenon of vowel lifting. Lastly, the varieties of PB from São Paulo and Rio Grande do Sul are, so far, the most similar to PL.

In *Aspects of Intonation in Speech of Schizophrenic Patients*, written by Waldemar Ferreira Netto (University of São Paulo), Marcus Vinícius Moreira Martins (State University of Minas Gerais) and Ana Cristina Aparecida Jorge (University of São Paulo), a methodology for characterizing intonation in the speech of subjects with schizophrenia is developed, which can facilitate the diagnosis of patients. For this study, the authors carried out an experiment with 10 patients with schizophrenia and 10 control subjects, using the ExProsodia application and Shannon's entropy calculation for the analyzed sentences. The results obtained indicate that the speech intonation of subjects with schizophrenia shows a strong tendency to decrease the amount of information that can be detected through the variation of F0.

In *EJA and the Elderly: Paths to Inclusion*, Maria Cecília de Magalhães Mollica (from the Federal University of Rio de Janeiro) proposes a review of the book *Alfabetização e idosos – um tema urgente no Brasil* (Literacy and the Elderly – an Urgent Topic in Brazil), by Kátia Abreu (from the State University of Rio de Janeiro) and Maria Carlota Rosa (from the Federal University of Rio de Janeiro). By summarizing aspects of its 18 chapters linked to writing systems, learning to read and literacy through Youth and Adult Education, the researcher invites the academic-professional community in the area of Letters and Linguistics to read the book, specifically those interested in education movements and routes in Brazil and in the link between the theoretical-scientific world and public policies.

This dossier is the result of the gathering of these abundant contributions. They offer the reader a valuable panoramic view in Linguistics, based on the variable uses of the Portuguese and Occitan languages as well as on a wide repertoire of theoretical, methodological, analytical and descriptive perspectives of the phenomenon of variation observed in different academic spaces in Brazil and other American and European countries.

The organizers of this dossier are very grateful to the journal's editorial team for the opportunity, and to the issue's design team. We are very grateful to the advisory board and the reviewers who gave their time and effort to the examination of each article submitted, in order to select the articles found here. We are also grateful to the professors, students and other professionals of the Postgraduate Program in Vernacular Letters for their support in the wide dissemination of this dossier and of the wonderful issue of which it is part. We also thank the team from Paul-Valéry University Montpellier III, for the fruitful dialogue since the conception of the VariaR Project. Finally, we invite the reader to appreciate the dossier, in the certainty that it contains ample matter for reflection, description and/or action that may impact other science and education movements and spaces. We hope, therefore, that the reader will benefit from it.

References

AYRES-BENNETT, Wendy; CARRUTHERS, Janice (ed.). *Manual of Romance Sociolinguistics*. Berlin: De Gruyter, 2018, xi + 793 pp.

BRANDÃO, Silvia F. (ed.). *Duas variedades africanas do Português: variáveis fonético-fonológicas e morfossintáticas*. São Paulo: Blucher Open Access, 2018.

DE PAULA, Alessandra; GOMES, Danielle K.; SILVEIRA, Eliete F. B.; MACHADO VIEIRA, Marcia dos S; VIEIRA, Silvia R. (ed.). *Uma História de Investigações sobre a Língua Portuguesa: Homenagem a Silvia Brandão*. São Paulo: Blucher Open Access. 2018. <https://www.blucher.com.br/livro/detalhes/uma-historia-de-investigacoes-sobre-a-lingua-portuguesa-homenagem-a-silvia-brandao-1406>

LEDGEWAY, Adam; MAIDEN, Martin (ed.). *The Oxford guide to the romance languages*. Oxford; New York: Oxford University Press, 2016.

LÓPEZ, Laura Álvarez; GONÇALVES, Perpétua; AVELAR, Juanito Ornelas de (ed.). *The Portuguese language continuum in Africa and Brazil*. Amsterdam: John Benjamins Publishing Company, 2018.

MACHADO VIEIRA, Marcia dos Santos; MEIRELES, Vanessa (ed.). *Variação em português e em outras Línguas Românicas*. São Paulo: Blucher Open Access, 2022. <https://www.blucher.com.br/variacao-em-portugues-e-em-outras-linguas-romanicas>

MARTINS, Marco Antonio; ABRAÇADO, Jussara. (ed.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015.



ANÁLISE DE DADOS LINGUÍSTICOS Entrevista com *Adriana Picoral*

LINGUISTIC DATA ANALYSIS Interview with *Adriana Picoral*

*Adriana Picoral*¹

*Marcia dos Santos Machado Vieira*²,

Vanessa Meireles

Ravena Beatriz de Sousa Teixeira

Mariana Gonçalves da Costa

RESUMO

Adriana Picoral é Professora Assistente de prática no Departamento de Ciência da Computação e membro do corpo docente afiliado no Programa de Pós-Graduação Interdisciplinar de Aquisição e Ensino de Segunda Língua da Universidade do Arizona. Ela é a fundadora da R-Ladies Tucson, que faz parte de uma organização mundial para promover a diversidade de gênero na comunidade R. É doutora em Linguística Aplicada e Bacharel em Ciência da Computação (UFRGS). Sua pesquisa baseia-se em Linguística Computacional e Linguística de Corpus, no intuito de esclarecer o uso, a aquisição e o desenvolvimento multilinguísticos. A presente entrevista focaliza temas relativos à ciência da análise de dados linguísticos, ciência da informação e ciência computacional.

PALAVRAS-CHAVE: Sociolinguística; Linguística de Corpus; Linguística Computacional; Processamento e análise de dados linguísticos; Português e Inglês.

1 A pesquisadora entrevistada, Professora Doutora *Adriana Picoral* (adrianaps@arizona.edu), atua na Universidade do Arizona, na área de coleta, processamento, manutenção e análise de dados linguísticos em ambientes acadêmicos e além.

2 *Marcia dos Santos Machado Vieira* (marcia@letras.ufrj.br, docente da UFRJ, Pesquisadora do CNPq e Cientista do Nosso Estado/Faperj), *Vanessa Meireles* (vanessa.meireles@univ-montp3.fr, docente da UPVM), *Ravena Beatriz de Sousa Teixeira* (ravena_beatriz@letras.ufrj.br, ravenabst@gmail.com, discente da UFRJ) e *Mariana Gonçalves da Costa* (marianag.costta@gmail.com, discente da UFRJ) – as entrevistadoras – são pesquisadoras vinculadas a projetos de pesquisa que têm articulação interinstitucional: Programa de Pós-Graduação *stricto sensu* em Letras Vernáculas (Língua Portuguesa) da Universidade Federal do Rio de Janeiro (UFRJ) e RESO (Recherches sur les suds e les orientes) da Universidade Paul-Valéry/Montpellier 3. As três primeiras são membros do Projeto VariaR (Variação em línguas românicas). E Marcia Machado Vieira, Ravena Teixeira e Mariana Costa são pesquisadoras do Projeto PREDICAR (Formação e expressão de predicados complexos e predicções: estabilidade, variação e mudança construcional).

ABSTRACT

Adriana Picoral is an Assistant Professor of practice in the Computer Science Department, and an affiliated faculty member in the interdisciplinary Graduate Program of Second Language Acquisition and Teaching at the University of Arizona. She is the founder of the R-Ladies Tucson, which is part of a world-wide organization to promote gender diversity in the R community. She holds a PhD in applied linguistics and a bachelor's degree in computer science (UFRGS). Her research draws from Corpus Linguistics and Computational Linguistics to shed light on multilingual language use, acquisition, and development. This interview focuses on topics related to the science of linguistic data analysis, information science and computational science.

KEYWORDS: Sociolinguistics; Corpus Linguistics; Computational Linguistics; Processing and analysing language data; Portuguese and English.

Questão 1

Iniciamos agradecendo à Profa. Dra. Adriana Picoral por aceitar nosso convite para esta entrevista e pedindo-lhe que nos conte um pouco sobre sua trajetória acadêmica e sobre como se iniciou seu interesse por análise linguística e por ciência da informação. O que destacaria no seu percurso acadêmico-científico? Quais são seus projetos atuais? Qual é sua área e sua atuação na Universidade do Arizona?

A minha trajetória acadêmica é interdisciplinar e variada, como minha atual área de atuação. Minha graduação foi em Ciência da Computação na UFRGS, e minha bolsa de iniciação científica na época foi com a Dr. Margarete Schlatter, Professora de Prática de Ensino de Línguas Adicionais e de Linguística Aplicada. No meu doutorado em linguística aplicada eu continuei esse processo de combinar minha formação em métodos computacionais com análise linguística no meu trabalho com as professoras Dras. Ana Carvalho, em Sociolinguística, e Shelley Staples, em Linguística de Corpus. Meus projetos atuais incluem o estudo de variação linguística do Português Brasileiro no Twitter e do Inglês acadêmico por aprendizes. Minha área de atuação na Universidade do Arizona é Ciência de Dados em termos amplos, e, em termos mais específicos, é análise quantitativa de dados de dados linguísticos.

Questão 2

Poderia nos apresentar um pouco das metas e dos desafios perspectivados pela equipe de pesquisa que coordena? Quais são os perfis dos subprojetos de pesquisa, das áreas de contribuições destes? Em que medida e de que maneira a ciência de dados, informações, textos se relaciona à ciência linguística?

O maior desafio atual relacionado à análise quantitativa de dados linguísticos é estabelecer quais métodos estatísticos são mais úteis para a generalização de resultados de análises de variação linguística em larga escala. Os subprojetos de pesquisa que coordeno incorporam métodos originários da Sociolinguística de Variação e da Linguística de Corpus, dando atenção maior a como minerar e analisar grandes quantidades de produção natural de linguagem. A

Sociolinguística de Variação historicamente faz uso de métodos estatísticos detalhados para a análise do que, hoje em dia, é considerado uma quantidade de dados menor e mais selecionados. Por outro lado, a Linguística de Corpus foca na coleta de muitos dados (em milhões e até bilhões de palavras), com métodos de análise estatísticas desses dados ainda não estabelecida pelos pesquisadores na área. Não há um conjunto de testes estatísticos usados na Linguística de Corpus, alguns pesquisadores reportam só porcentagens, ou frequências normalizadas, enquanto outros aplicam regressão para diferentes propósitos. Para dar um exemplo mais concreto, o estudo de variação do sujeito pronominal no Português Brasileiro é muito produtivo na área de Sociolinguística, com uma série de estudos pequenos (com dezenas de participantes e centenas de tokens) com dados coletados em diversas regiões do Brasil. Nosso grupo de pesquisa construiu um corpus de tweets em Português (produzidos por centenas de usuários, e milhares de tokens) que permitiu o estudo da variação do sujeito pronominal no Português Brasileiro em diversas regiões do Brasil no mesmo meio de comunicação, no mesmo período de tempo (Picoral et. al, 2021). Implementamos diversos métodos e ferramentas computacionais para minerar dados para análise estatística confiável e generalizável, a partir dos dados crus extraídos do Twitter. O próximo passo para nosso grupo é organizar e compartilhar essas ferramentas e métodos com outros pesquisadores com o propósito de expandir esse tipo de pesquisa em larga escala.

Questão 3

Desde 2019, você atua como embaixadora da iniciativa global *Women in Data Science*. Como se configura o projeto? O que a motivou a participar da proposta? Segundo sua perspectiva, como este contribui para o posicionamento da figura feminina nos espaços acadêmico-científicos?

Representação feminina em específico, e diversidade em termos mais gerais, são problemas em espaços relacionados à informática, tanto na esfera acadêmica quanto na indústria. Na minha graduação na UFRGS nós mulheres tínhamos menos de 10% de representatividade. Ser embaixadora da iniciativa global *Women in Data Science* faz parte do meu trabalho de inclusão de mulheres em um ambiente técnico que não recebe mulheres de braços abertos. O treinamento em métodos computacionais exige uma comunidade de prática que permite o processo de falhar e aprender com erros, e essa comunidade de prática não existe para mulheres em ambientes dominados por homens. Então se faz preciso criar ambientes dominados por mulheres, para dar espaço para mulheres, e outras pessoas excluídas dessas áreas técnicas, receberem esse treinamento que é tão essencial para avançar o conhecimento e a pesquisa em áreas que estão começando a utilizar mais e mais ferramentas computacionais, como a ciência linguística.

Questão 4

A partir de seu conhecimento empírico, poderia comentar brevemente sobre a relevância de saberes referentes ao tratamento, processamento e análise quantitativa de dados no contexto atual dos estudos linguísticos? Quais são os pontos que carecem de aprimoramento na área?

Acredito que o ponto mais importante que carece aprimoramento na área dos estudos linguísticos é reprodutibilidade, que possibilita a reprodução de todos os passos de tratamento, processamento e análise quantitativa de dados, a partir dos dados crus. Nosso forte nos estudos linguísticos é coletar dados, mas não temos métodos estabelecidos de o que fazer a partir desse dados, até mesmo pela ampla variedade de subcampos com diferentes perspectivas de análise linguística. Falta também a prática de ciência aberta nessa área no Brasil, seguindo iniciativas como a da Open Science Framework (<https://osf.io/>) que possibilitam o compartilhamento aberto de dados e métodos. A estatística não é uma ciência que é aplicada descasada dos outros métodos e conhecimentos de uma área. Pelo contrário, é preciso entender bem quais as perguntas e perspectivas que são exploradas em estudos linguísticos para entender como a estatística pode responder essas perguntas em termos quantitativos. A partir do compartilhamento aberto dos métodos que estão sendo atualmente explorados e utilizados na linguística é que podemos construir um corpo de conhecimento sobre os métodos estatísticos mais eficazes, com reprodução desses métodos de maneira sistemática em diferentes estudos.

Questão 5

Em sua opinião, o tópico (tratamento de dados na análise linguística) é abordado de forma satisfatória no que concerne à formação de linguistas em território brasileiro? Considerando sua formação (inclusive, no Brasil), em que falta aos currículos dos cursos de graduação (Bacharelado e Licenciatura) na área de Letras e Linguística investir, para o enriquecimento das análises linguísticas que consubstanciam o que é descrito?

O tratamento de dados na análise linguística não é abordado de forma satisfatória no que concerne à formação de linguistas nem em território brasileiro (nem nos Estados Unidos). Como aluna de graduação na UFRGS há duas décadas, eu tive que lutar muito para ter uma educação interdisciplinar. Apesar de algumas colaborações existirem, a conversa entre áreas como Ciência da Computação e Letras e Linguística em universidades Brasileiras, por exemplo, é muito limitada. Sem se falar na inclusão de cursos de estatística que oferecem uma perspectiva aplicada de conceitos estatísticos em práticas de pesquisa que envolvem análises linguísticas. Ainda existe muita resistência discente e docente em atravessar essa fronteira fictícia entre ciências humanas e ciências exatas. Mas cultura não muda sem exemplos, e eu acredito que, a partir de mais grupos e projetos de pesquisa que incluem métodos quantitativos de análise

linguísticas, maior nossa comunidade de prática e mais pessoas qualificadas para ministrar cursos aplicados de estatística e métodos computacionais em cursos de graduação na área de Letras e Linguística no Brasil. Essa lacuna vai demorar a ser preenchida, por isso a necessidade de fomentar ainda mais e com mais emergência esse tipo de pesquisa.

Questão 6

Durante seu minicurso ministrado no I Congresso PREDICAR³, houve uma discussão quanto à falta de fundamentação teórica na aplicação de métodos de análise estatística nas pesquisas linguísticas. Quais seriam as possíveis repercussões desse problema na leitura e adequação da análise? Quais propostas podem ser (ou estão sendo) feitas para solucionar essa questão na área da linguística?

As ferramentas computacionais que usamos hoje em dia para análise estatística e a quantidade de tutoriais e informação gratuita na internet facilitam muito a aplicação desses métodos. No entanto, não existe transparência em como esses métodos são aplicados na maior parte dos meios de publicação científica na área da linguística. Então, alunos, professores, e pesquisadores que estão na posição de começar a aprender e utilizar estatística não tem referências de como interpretar os resultados estatísticos, só existe referência de como rodar certos testes na sintaxe em linguagens de programação como Python e R. Porém, usar um certo tipo de função em uma dessas linguagens de programação não é suficiente. A análise de dados de maneira quantitativa se faz confiável a partir do tratamento de dados até a interpretação de resultados. Rodar uma função para um teste estatístico é uma pequena parte do que é necessário para produzir resultados de uma análise linguística. A maior repercussão dessa falta de entendimento sobre métodos estatísticos é a carência de reprodutibilidade. Temos projetos com alta qualidade de coleta de dados, mas que falham na análise estatística, não produzindo os resultados necessários que adicionam à literatura da área. Acredito que grupos como PREDICAR e o programa LAEAL da PUCSP são os protagonistas atuais que procuram avançar a formação de alunos nesses métodos computacionais e estatísticos na área da linguística no Brasil. Ainda mais grupos, mais projetos de pesquisa, que se preocupem com a qualidade das interpretações estatísticas de resultados são necessários para acelerar essa questão.

Questão 7

Considerando o seu trabalho com dados de uso da rede social *Twitter*, quais desafios e potencialidades foram encontrados por você e pela sua equipe quanto aos dados oriundos dessa plataforma? Quais especificidades estariam envolvidas na coleta e no tratamento de dados virtuais?

3 Disponível em: <https://www.youtube.com/watch?v=fe1DIiIILxE&list=PLY1bzUR7D2so-Xo9yP0yOhbPuu2el2YwtF>. Acesso em: 31 de maio, 2022.

As redes sociais têm relevância para a nossa comunicação diária, muita linguagem ocorre exclusivamente através desses meios. No entanto, pouca pesquisa e entendimento existe sobre esse tipo de linguagem, especialmente em diferentes variedades do Português e outras línguas menos pesquisadas. Justamente por esse carência de estudos é que as potencialidades são muitas. O maior desafio com dados coletados do Twitter é a falta de informação sobre a maioria dos usuários. Não temos acesso a metadados demográficos como faixa etária, etnia, e classe social. Claro, se a coleta se faz a partir de perfis de celebridades e outras figuras públicas, mais se sabe sobre o suposto interlocutor. Outro desafio é a variedade ortográfica. Por exemplo, no nosso estudo pronominal, o sujeito *nós* produz grafias como *nóis*, *nos*, *nóix*, e assim por diante. Poucas ferramentas existem para a normalização ortográfica em Português, então o tratamento e anotação de dados exige ainda muito processamento manual, o que é dispendioso em termos de tempo e esforço. O maior potencial é com certeza a larga escala de coletas de dados, que apesar de terem que ser parcialmente processados manualmente, ainda demonstra vantagens em termos de custo comparado à transcrição de comunicação oral.

Questão 8

Quais são geralmente os requisitos de uma análise linguística de perfil científico considerada rigorosa e confiável? Em que as tecnologias e ciências da informação e da computação podem auxiliar concretamente? Que recomendações ressaltaria nesse sentido?

Uma análise linguística de perfil científico rigoroso e confiável é aquela que é reproduzível. Nossa capacidade tecnológica de armazenamento de dados é extensa, o que permite que todos os estágios de processamento dos dados sejam armazenados, dos dados crus, até a versão final processada dos dados. Tutoriais e recursos de como organizar projetos colaborativos que são reproduzíveis existem de graça e fácil acesso na internet. Plataformas como a Open Science Framework (osf.io) contém exemplos e modelos a partir do primeiro passo em qualquer processo, o pré-registro de estudos, até a comunicação final de resultados. A minha maior recomendação é usar outros projetos abertos como base de organização e gerenciamento de projetos, e ser transparente em relação a todos os passos em qualquer projeto de pesquisa.

Questão 9

Em que medida esses requisitos repercutem em ações desenvolvidas em prol da constituição ou reunião de acervos de dados linguísticos (grandes bancos de dados), a fim de potencializar diversificadas análises linguísticas? Tendo em vista movimentos como, por exemplo, o de uma equipe de comissões científicas e estratégicas da Abralin e do GT de Sociolinguística da ANPOLL trabalhando em prol de um repositório digital intitulado *Projeto Plataforma da Diversidade Linguística Brasileira*, o que, a partir de sua ótica de formação em ciência da informação, seria estratégico considerar no planejamento das ações?

Grandes bancos de dados que são representativos da fala de comunidades bem definidas, como é o caso do Projeto Plataforma da Diversidade Linguística Brasileira, são essenciais para a diversificação de análises linguísticas. A estratégia de usar times grandes e diversos, como é o caso nesse projeto, é a melhor forma de garantir que o acervo seja de utilidade para outros grupos de pesquisa diversificados no Brasil e no exterior. Com os dados disponíveis a pesquisadores, um sistema de pré-registro de estudos se faz útil para não só a comunicação de projetos que estão em andamento, mas também para a abertura de colaborações interdisciplinares.

Questão 10

Que possibilidades de análises, recursos e perfis têm o potencial de enriquecer o tratamento de dados de variação linguística? Em que metanálises agregam poder exploratório a fenômenos em variação? Que articulação hoje em dia é possível perspectivar entre Sociolinguística e Ciência da Informação, ou a partir de Sociolinguística?

A Sociolinguística, especialmente a área de variação e mudança, tem seus métodos de análise quantitativo bem estabelecidos, com o uso de regressão logística com soma de contrastes desde os anos 70. Há muitos estudos em Sociolinguística que analisam as mesmas variáveis. Esses resultados agregados nos dão uma base de exploração para novos estudos mais inovativos, como os que fazem uso de dados de redes sociais. A Ciência da Informação entra com esse aspecto inovativo, que nos permite minerar e processar uma larga escala de dados. Esse aspecto inovativo casado com os métodos de análise quantitativa usados na Sociolinguística ajuda a expandir o que entendemos sobre essas variáveis sociolinguísticas.

Questão 11

Tendo em vista sua expertise em diversas ações formativas (minicursos, entre outras), o que geralmente tem tido maior acolhida entre pesquisadores (docentes e discentes) e o que normalmente lhes tem causado algum desconforto ou até os tem silenciado? Que (inter)ações, leituras e/ou atividades de formação lhes indicaria, no intuito de começarem a ver com outros olhos e possivelmente encarar com mais segurança o espaço (tecnológico) de análise de dados?

Existe um desejo e empenho entre os pesquisadores brasileiros por aprender novos métodos e ferramentas computacionais. Durante minhas ações formativas eu encontrei muito mais acolhimento do que o contrário. Os pesquisadores brasileiros, na minha experiência, já vão atrás de conhecimento e estão sempre se aprimorando, nunca param de estudar. O maior desconforto é a ideia de ciência aberta em todos os estágios de tratamento e análise de dados. É desconfortável saber que outros pesquisadores vão poder inspecionar e criticar o nosso trabalho por inteiro, incluindo o processo, e não só o produto final. Mas volto a iterar que grupos grandes, com colaborações interdisciplinares, que quebram as barreiras entre ciência humanas e exatas é

como vencemos esse desafios e criamos uma comunidade de prática capaz de treinar pesquisadores novatos e aprimorar pesquisadores experientes. A criação de grupos de leitura que abordam títulos disponíveis gratuitamente como R for Data Science (<https://r4ds.had.co.nz/>) e Introduction to Modern Statistics (<https://openintro-ims.netlify.app/>) são as minhas recomendações atuais.

Referências

PICORAL, A. MINICURSO – Analisando construções intensificadoras encontradas em tweets por meio do R. Youtube, 2022. Disponível em: <https://www.youtube.com/watch?v=fe1DIiLLxE&list=PLYIbzUR7D2soXo9yP0yOhbPuu2el2YwtF>. Acesso em 31 de maio, 2022.

PROJETO PLATAFORMA DA DIVERSIDADE LINGUÍSTICA BRASILEIRA. Museu da Língua Portuguesa, 2022. Disponível em: <https://www.museudalinguaportuguesa.org.br/projeto-plataforma-diversidade-linguistica-brasileira/>. Acesso em: 31 de maio, 2022.

PICORAL, A.; Stumpf, E.; GOULART, L.; DE BARROS, I.C.; SOMMER-FARIAS, B.; Matte, M.L.; GARCIA, M.C.; BERTHO, M.C.. Parallels Between Spoken and CMC Language: Do Tweets Reflect Spoken Language Choices?. In: *Proceedings of the 8th Conference on Computer-Mediated Communication (CMC) and Social Media Corpora (CMC-Corpora 2021)*, 28-29 October, 2021 (p. 84). <https://cmc-corpora.org/publications/cmc-corpora-2021/>



**PLACE SOCIALE ET VARIATION
DIASYSTÉMIQUE DE L'OCCITAN**

Interview avec *HERVÉ LIEUTARD*

**LUGAR SOCIAL E VARIAÇÃO
DIASSISTÊMICA DO OCCITANO**

Entrevista com *Hervé Lieutard*

Hervé Lieutard

*Marcia dos Santos Machado Vieira*¹

Vanessa Meireles

RÉSUMÉ

Hervé Lieutard est professeur titulaire de linguistique occitane à l'Université Paul-Valéry – Montpellier 3 depuis 2001. Ses recherches portent sur l'analyse de la variation dialectale en occitan dans une perspective phonologique mais également sur la place sociale de l'occitan à travers toute son histoire, de l'époque médiévale à l'époque contemporaine, étudiée à travers la vaste documentation écrite disponible dans cette langue. Dans cette interview, il fait un bref panorama des difficultés de la recherche en occitan en lien avec le déclin social de l'occitan, mais montre également comment les processus de revitalisation ont contribué à changer l'image de cette langue et à développer à nouveau sa transmission et sa socialisation.

MOTS-CLES: Occitan ; Langue régionale ; Diglossie ; Sociolinguistique ; Phonologie.

1 *Marcia dos Santos Machado Vieira* (marcia@letras.ufrj.br) est professeur-chercheuse à la Faculté des Lettres de l'Université Fédérale de Rio de Janeiro, affiliée au programme de troisième cycle en langues vernaculaires (langue portugaise) de la même université, et, parmi d'autres responsabilités, coordonne le projet franco-brésilien VariaR – Variation dans les langues romanes – en partenariat avec Vanessa Meireles (vanessa.meireles@univ-montp3.fr), maître de conférences en langue, culture et littérature du Brésil à l'Université Paul Valéry – Montpellier 3 (France) depuis 2017, membre de l'équipe de *Recherches sur les Suds et les Orient*s (ReSO).

Question 1

Professeur Hervé Lieutard, pouvez-vous nous parler brièvement de votre profil d'intérêts de recherche et de l'articulation de votre/vos recherche(s)?

La langue occitane est au centre de mes recherches depuis plus de vingt ans. La démarche que j'adopte pour l'étude de cette langue régionale de France combine à la fois une approche phonologique de la variation dialectale propre à cette langue et une approche sociolinguistique qui permet de mieux comprendre les effets du contact de langue, notamment entre occitan et français, sur les changements qu'a connus cette langue au fil de son histoire et sur l'évolution de ses usages et de sa place dans la société française.

L'occitan, malgré son développement particulièrement important au Moyen Âge et son prestige littéraire, n'est jamais devenu la langue officielle d'un État. En l'absence de forme officielle standard, c'est donc une langue pluridialectale qui se caractérise par une variation linguistique importante d'un côté à l'autre de son espace (elle recouvre un bon tiers de l'espace français, soit 33 départements², mais elle est aussi présente dans quatorze vallées du Piémont italien et dans le Val d'Aran en Espagne). Au-delà de la variation dialectale observable dans cette langue, les structures communes aux divers dialectes permettent de définir un diasystème occitan dans lequel l'intercompréhension fixe les limites externes entre les six grandes variétés qui composent cette langue (languedocien, provençal, gascon, limousin, auvergnat et vivaro-alpin). L'approche phonologique que je développe dans mes recherches se nourrit des données issues des acquis de la dialectologie de terrain. Depuis plusieurs années, j'ai tenté de montrer que l'utilisation des théories phonologiques sur la syllabe en constituants organisés permettait non seulement de rendre compte de la structure interne de cette langue, mais aussi de sa variation dialectale, notamment au niveau des finales syllabiques. Cette approche permet notamment de dépasser l'approche géolinguistique traditionnelle de la variation basée sur des isoglosses. La typologie syllabique que j'ai définie pour l'occitan sert d'unité de mesure de la variation dialectale tout en permettant de rendre compte de fonctionnements supradialectaux qui rendent compte de l'unité sous-jacente de cette langue au-delà de l'évolution inégale de ses dialectes dans le temps et dans l'espace. La diversité dialectale de cette langue représente aujourd'hui encore un vaste terrain d'investigation pour les phonologues. Ces dernières années, j'ai également élargi mon champ de recherche en m'intéressant à la typologie syllabique propre aux langues romanes, avec une attention particulière portée au catalan, une langue qui s'est séparée peu à peu de l'occitan à partir du Moyen Âge pour des raisons historiques, mais qui reste totalement compréhensible pour des locuteurs occitanophones.

Un autre aspect de mes recherches porte sur l'émergence des standards linguistiques écrits à l'époque médiévale pour lesquels l'occitan représente sans doute un des modèles incomparables dans le domaine roman. L'occitan ayant été pionnier dans la mise en place de

2 39 si l'on compte les départements dans lesquels il n'est que partiellement présent.

pratiques d'écritures standardisées au Moyen Âge, à la fois dans le registre lyrique des Trobadors, mais aussi pour les usages pragmatiques de la langue (traités rhétoriques, grammaires, textes juridiques, notariaux, administratifs ou scientifiques), je travaille également sur l'édition numérique de textes occitans anciens dans des formats normalisés (XML-TEI) qui permettent aujourd'hui d'envisager une interopérabilité entre les divers documents édités dans ce format dans le but de proposer un traitement partiellement automatisé des données obtenues (analyse morpho-syntaxique, lexicale, mais aussi graphématique). Ces travaux permettent d'évaluer les diverses évolutions qui caractérisent l'occitan de l'époque médiévale à nos jours et en particulier d'évaluer et de mesurer la continuité des usages lexicaux entre le Moyen Âge et l'époque contemporaine.

L'approche sociolinguistique qui caractérise mes recherches est liée au fait qu'il est impossible de travailler sur l'occitan sans s'intéresser également à la place sociale de cette langue et aux effets du contact linguistique sur ses usages. La revue *Lengas*³ que je dirige depuis plusieurs années est une revue de sociolinguistique qui donne une place importante à cette question du contact linguistique, pour l'occitan, mais aussi pour toutes les autres langues en situation de diglossie. Ces travaux s'inscrivent dans le prolongement des modèles théoriques développés par l'école sociolinguistique montpelliéraine autour de Robert Lafont qui, en collaboration avec les sociolinguistes catalans, est à l'origine d'une analyse originale de la situation des langues dominées ou des langues en contact, notamment à travers une relecture critique du concept de diglossie proposé par Ferguson (1959) et Fishman (1967). Cette perspective sociolinguistique a guidé mes recherches sur la question de l'écrit occitan du Moyen Âge à nos jours, que ce soit pour l'étude de la littérature occitane ou des usages écrits non-littéraires. Ces dernières années, j'ai aussi consacré un certain nombre d'études aux usages graphiques en occitan à d'autres époques, dans la mesure où ils sont en quelque sorte le miroir de la place de l'occitan dans la société française, notamment dans la période qui, du XVe au XIXe siècle, se caractérise par un déclassement social de cette langue.

Question 2

Quel est le statut linguistique et social de l'occitan : langue maternelle, langue seconde? De combien de locuteurs est-ce la langue, de quelle(s) communauté(s), de quels types de textes oraux et écrits (de documents et/ou supports officiels, par exemple)?

L'occitan a longtemps souffert d'un déficit de visibilité dans la société française à cause des préjugés qui entouraient sa pratique et d'une politique linguistique qui, depuis la Révolution Française notamment, a tenté d'éradiquer les « patois » et de généraliser l'usage du français dans les pratiques orales. En France les représentations concernant les langues ont été et sont sans doute aujourd'hui encore largement formatées par le modèle de la langue française qui s'est imposé institutionnellement et qui a conduit à confondre la hiérarchie administrative

3 <https://journals.openedition.org/lengas/>.

officielle des langues (langue nationale et langue régionale) avec une hiérarchie de qualité de ces langues. Ce n'est pas le lieu de définir ici le long processus de minoration de l'occitan dans la société française, mais il est bon de rappeler seulement que la construction du français officiel, en particulier depuis la Révolution, s'est développé à travers un discours de négation de l'altérité linguistique qui a consisté notamment à déconstruire idéologiquement tout l'espace culturel et linguistique de l'occitan. Même si l'occitan a été une langue médiévale prestigieuse, son usage écrit administratif a commencé de se réduire au XVI^e siècle avec l'officialisation du français écrit (Ordonnance de Villers-Cotterêts en 1539, sous le règne de François Ier). Suite aux politiques linguistiques menées par l'État français pour la promotion du français oral, en particulier depuis la Révolution française, la transmission familiale a peu à peu baissé, mais c'est surtout à partir de la mise en place de l'école obligatoire à la fin du XIX^e siècle (Lois Jules Ferry, 1880-1881) que la connaissance du français par tous est devenue obligatoire. Les mouvements renaissantistes et littéraires occitans apparus au XIX^e siècle (Félibrige, occitanisme...) n'ont pas été en mesure d'inverser cette tendance à tel point que la transmission familiale de l'occitan s'est arrêtée vers le milieu du XX^e siècle.

Parmi les diverses langues de France, il est aujourd'hui plus facile d'identifier le breton, le corse, le basque voire même le catalan parce que ces langues portent aussi le nom d'une région alors que l'occitan qui ne s'est jamais constitué en entité politique distincte ne porte que le nom de sa langue (occitan ou langue d'oc vient de la façon dont on dit « oui » dans cette langue, à savoir « òc »).

Il est difficile aujourd'hui d'estimer précisément le nombre de locuteurs de l'occitan dans la mesure où la construction de l'identité française autour de l'idée d'un peuple indivisible ne permet pas que dans les statistiques réalisées en France au niveau national on puisse poser des questions qui pourraient faire penser qu'il existe des identités ou des groupes ethniques distincts. Au début du XX^e siècle le linguiste Jules Ronjat estimait à 10 millions le nombre de locuteurs occitanophones. L'enquête réalisée par l'Ofici public de la lenga occitana (OPLLO) en 2020 dans deux des quatre grandes régions occitanophones (Occitanie et Nouvelle-Aquitaine) nous apprend que le nombre de locuteurs qui déclare parler occitan sans difficulté s'élève environ à 7% de la population de ces territoires⁴. Dans la mesure où la transmission familiale s'est interrompue progressivement dans les décennies qui ont suivi la mise en place de l'école obligatoire, les derniers locuteurs qui ont hérité de l'occitan comme langue maternelle sont très âgés. La transmission de l'occitan ne se fait plus que comme langue seconde, majoritairement dans le cadre scolaire.

Quoiqu'il en soit, la langue occitane dispose d'une grande richesse documentaire qui couvre toutes les périodes, du Moyen Âge à nos jours. Si les chansonniers des Trobadors ont fait l'objet de nombreux travaux de recherche en France et à l'étranger, très peu de travaux

4 <https://www.ofici-occitan.eu/fr/restitution-des-resultats-de-lenquete-sociolinguistique/>.

ont été consacrés à l'écriture pragmatique, administrative ou scientifique en occitan à cette époque bien que les documents soient très abondants. Les travaux de numérisation et d'éditions numériques réalisés en France ou en Europe permettent de plus en plus en plus de proposer un accès ouvert à des documents divers qui vont du Moyen âge (Dictionnaire informatique de l'occitan médiéval⁵, textes littéraires⁶, textes d'occitan pragmatique⁷, etc.) à nos jours (<https://occitanica.eu/enciclopedia>). En dépit d'un statut diglossique qui a fortement limité le développement de cette langue, l'usage oral et littéraire de l'occitan ne s'est jamais interrompu jusqu'à aujourd'hui. Un mouvement renaissantiste littéraire occitan s'est amorcé au XVIe siècle, puis à nouveau au XIXe siècle avec la fondation du Félibrige⁸ autour du poète Frédéric Mistral qui a aussi rédigé le grand dictionnaire à portée encyclopédique et ethnographique qui fait toujours référence aujourd'hui, *Lou Tresor dóu Felibrige*⁹. Il existe aujourd'hui encore une importante production littéraire en occitan et plusieurs maisons d'édition qui publient des ouvrages en occitan (prose¹⁰, poésie¹¹...). Il existe également plusieurs radios occitanes¹², quelques journaux d'information, notamment en format électronique¹³ et quelques émissions en occitan sur les chaînes de télévision régionales¹⁴.

D'un point de vue lexicographique, les premiers dictionnaires occitans ont commencé à apparaître dès le XVIIIe siècle. Les travaux de recueil linguistique ont commencé dès le XIXe siècle avec la réalisation des premiers atlas et des premiers travaux de dialectologie. Des enquêtes se poursuivent encore aujourd'hui auprès des derniers locuteurs natifs. Les diverses campagnes d'enquêtes ont conduit à mailler tout le territoire occitanophone. Les données de terrain recueillies dans les divers atlas réalisés entre le XIXe et le XXIe siècle sont aujourd'hui rassemblées dans le Thesoc¹⁵, une base informatique développée par une équipe de chercheurs de Nice regroupant toutes les données dialectales de l'occitan. La documentation orale en occitan est importante, mais ces données sont éparpillées dans divers centres¹⁶. Certaines ont cependant pu faire l'objet d'un traitement récent concernant l'analyse de la microvariation syntaxique dans les langues de France¹⁷. Quoiqu'il en soit, le centre le plus important pour la

5 <http://www.dom-en-ligne.de>.

6 Par exemple, le répertoire informatisé de la littérature troubadouresque, Rialto (<http://www.rialto.unina.it/>).

7 Par exemple, l'édition récente des registres du Consulat médiéval de Montpellier (<http://thalamus.huma-num.fr>).

8 <http://www.felibrige.org/>.

9 <https://gallica.bnf.fr/ark:/12148/bpt6k74854.image>.

10 Par exemple, IEO-edicions, https://ideco-dif.com/ieo_edicions.

11 Par exemple, les Editions Jorn, <http://www.editions-jorn.com/catalogue.htm>.

12 Par exemple, Ràdio Lengadòc, <https://radiolengadoc.com>.

13 Par exemple, Jornalet, <https://www.jornalet.com>.

14 Par exemple Viure al País, <https://france3-regions.francetvinfo.fr/occitanie/emissions/viure-al-pais-0>.

15 <http://thesaurus.unice.fr/>.

16 Par exemple la Talvera, <http://cordae-talvera-documentation.kentika.fr>.

17 <http://symila.univ-tlse2.fr>.

documentation en occitan est Lo CIRDOC – Institut occitan de cultura¹⁸. Il s’agit d’un pôle associé à la Bibliothèque Nationale de France (BNF) qui a pour but de rassembler toutes les ressources et les savoirs du domaine occitan dans une visée encyclopédique, accessibles notamment à travers son portail multimédia Occitanica¹⁹.

Question 3

Comment le gouvernement français et la société française en général considèrent-ils la langue occitane, son utilisation et son enseignement? Quelles politiques gouvernementales nationales, politiques publiques, existent pour promouvoir et valoriser cette langue en France et au-delà?

Le statut officiel de l’occitan aujourd’hui est multiple dans la mesure où il est parlé dans trois États différents. En Espagne, dans le Val d’Aran, l’occitan est langue co-officielle avec le catalan et le castillan et à ce titre est utilisé officiellement dans l’administration, conformément au statut d’autonomie de la Catalogne, mais cette officialité de l’occitan en Catalogne ne concerne qu’une infime partie de l’ensemble des locuteurs de l’occitan puisqu’il y a moins de 10000 habitants dans le Val d’Aran. La proportion de l’enseignement en occitan dans cette région s’élève à 60% jusqu’à l’entrée au collège, mais diminue ensuite. En Italie, dans les vallées du Piémont, l’occitan a le statut de langue protégée par la loi, mais l’italien reste la seule langue officielle et l’occitan n’est pas enseigné dans les écoles. En France, le statut linguistique de l’occitan est simplement celui d’une langue régionale. L’article 75-1 de la Constitution française, voté en 2008, stipule que « les langues régionales font partie du patrimoine de la France », mais dans la mesure où seul le français est la langue officielle de la République (Article 2 de la Constitution : « la langue de la République est le français »), sa pratique se limite aux cercles privés (familles, amis...), même si son enseignement est autorisé de la maternelle à l’université depuis 1951 (loi Deixonne). Depuis cette époque, l’enseignement de l’occitan s’est toutefois largement développé dans l’enseignement public (cours de langue, classes bilingues), mais aussi dans le privé (enseignement immersif).

Les positions officielles du gouvernement français ne favorisent que dans une très faible mesure le développement de l’occitan à l’heure actuelle. La Charte européenne des langues régionales et minoritaires, élaborée par le Conseil de l’Europe afin d’assurer la protection des langues régionales historiques ou des langues des minorités de l’Europe a été signée par le gouvernement français en 1999, mais bien que la signature de la Charte se soit faite à partir d’une sélection de mesures modérées, le texte a finalement été rejeté par le Conseil constitutionnel français, car, selon ses déclarations « La Charte européenne des langues régionales ou minoritaires comporte des clauses contraires à la Constitution » et porte atteinte

¹⁸ <https://www.oc-cultura.eu>.

¹⁹ <https://occitanica.eu/>.

« aux principes d'indivisibilité de la République, d'égalité devant la loi et d'unicité du peuple français » (Décision No 99-412 DC du 15 juin 1999). En donnant des droits spécifiques à des communautés linguistiques, la Charte européenne irait donc selon le Conseil constitutionnel à l'encontre de l'article 1er de la Constitution française qui proclame l'indivisibilité de la République, mais elle serait également selon lui difficilement compatible avec l'article 2 qui proclame que « la langue de la République est le français ».

Malgré ces blocages qui empêchent une véritable reconnaissance officielle de l'occitan, il faut quand même souligner quelques avancées institutionnelles. Depuis 2016, une convention entre l'État français et les régions Occitanie et Nouvelle-Aquitaine a permis la mise en place de l'Ofici public de la lenga occitana (OPLLO)²⁰, un organisme qui contribue à valoriser l'occitan dans la société en soutenant les processus de formation des enseignants, des étudiants et le développement de l'occitan dans les structures scolaires. Il subventionne également des projets qui contribuent à la transmission et la socialisation de la langue occitane. Malheureusement cet organisme ne couvre que deux des quatre grandes régions dans lesquelles l'occitan est langue historique. Depuis 2011, existe également Lo Congrès²¹ qui propose des outils linguistiques en ligne (dictionnaires, conjugueur, banque de données terminologiques, numérisation des dictionnaires historiques, applications de synthèse et de reconnaissance vocale...). Je suis actuellement vice-président du Conseil linguistique de cet organisme de régulation linguistique soutenu par le Ministère de la Culture. Un projet de grand dictionnaire historique de la langue occitane est actuellement en cours d'élaboration dans le cadre des travaux de cet organisme.

Question 4

Comment voyez-vous l'insertion de cette langue dans l'ensemble des langues romanes ? Quelle attention en termes de pratiques de recherche, d'enseignement et de diffusion (y compris via la voie de la science ouverte) a-t-elle reçu dans les contextes universitaires et scientifiques centrés sur les langues romanes?

Au-delà du fait que l'occitan est une langue en danger de disparition selon l'Unesco, l'occitan fait encore partie aujourd'hui des « understudied languages », autrement dit des langues sous-étudiées pour lesquelles il reste d'importants travaux de recherche à faire pour mieux les connaître. La reconnaissance scientifique d'un espace linguistique occitan a mis beaucoup de temps à émerger et de ce fait l'occitan n'est devenu que très récemment un objet d'étude scientifique à part entière. Dans les dernières décennies du XIXe siècle, au moment où s'est développé un nouvel intérêt pour la littérature médiévale occitane des *Trobadors* et où ont commencé les premiers travaux scientifiques de linguistique sur l'occitan, plusieurs représentants de la linguistique française officielle, aux premiers rangs desquels, Gaston

20 <https://www.ofici-occitan.eu/fr/accueil/>.

21 <https://locongres.org/oc/>.

Paris et Paul Meyer, tous deux professeurs au Collège de France, ont largement développé un discours de négation et de rejet de la notion même d'espace linguistique occitan et de ses variétés dialectales pour mieux appuyer le projet d'unification nationale autour du français et de sa capitale, en considérant que les variétés d'occitan n'étaient rien d'autre que des variétés éloignées et « déformées » du français, ce que l'on appelait péjorativement des « patois ». Par cette forme de mépris, ils ont contribué à invalider et à déposséder les linguistes méridionaux de leur objet de travail et ont contribué à renforcer l'ensemble des préjugés véhiculés à cette époque sur les langues de France, privant par la même occasion les locuteurs occitanophones de la possibilité d'accéder à une forme de « conscience linguistique » occitane. À l'appui de ces théories idéologiques d'unification nationale s'est développée la notion de gallo-romania qui est en fait une application inappropriée de concepts historio-géographiques à la linguistique et qui montre surtout le poids de l'idéologie dans la construction de frontières linguistiques arbitraires. La reconstruction et la légitimation d'un espace linguistique occitan a nécessité en premier lieu un long travail de déconstruction — sans doute encore inachevé — des préjugés et des représentations à la fois endogènes et exogènes dont faisait l'objet cette langue.

Quoiqu'il en soit le principe de l'intercompréhension occitane est toujours demeuré à la base de toutes les études linguistiques et sociolinguistiques consacrées à l'occitan des XXe et XXIe siècles. C'est le linguiste Pierre Bec qui a adapté à l'occitan la notion de diasystème, proposant une nouvelle structuration de l'occitan qui dépassait les limites dialectales traditionnelles définies avant lui. Dans la dernière édition de son ouvrage de référence sur la langue occitane (BEC, 1995, 6), Pierre Bec a d'ailleurs élargi la notion de diasystème à tout l'espace qu'il a défini comme occitano-roman, incluant l'occitan et le catalan dans cet ensemble distinct du gallo-roman. Dans la perspective de Pierre Bec, le diasystème permettait de voir dans les dialectes de l'occitan et du catalan en synchronie non pas l'émergence de formes nouvelles de langue, mais plutôt l'évolution inégale dans le temps et l'espace d'un seul protosystème linguistique original (BEC, 1972, 46).

Quoiqu'il en soit, l'occitan occupe une place centrale dans l'ensemble des langues romanes et représente sans doute un des ponts les plus intéressants vers les autres langues romanes méridionales (italien, catalan, castillan, voire même portugais...). Malheureusement, en l'absence de forme officielle reconnue et d'une politique officielle de l'état français pour sa promotion, cette langue reste largement méconnue, aujourd'hui même sur son propre territoire, et en particulier par les linguistes romanistes. C'est surtout la littérature médiévale qui a suscité l'intérêt des chercheurs au niveau international, même si aujourd'hui de plus en plus de recherches concernent aujourd'hui la littérature moderne et contemporaine, la civilisation et la linguistique. Les chercheurs qui travaillent sur l'occitan en France sont peu nombreux, mais il existe heureusement un réseau international de chercheurs occitanistes regroupés au sein de

l'Association internationale d'études occitanes (AIEO)²², dont je suis le secrétaire général, qui rassemble environ 300 chercheurs et universitaires du Monde entier qui s'intéressent à l'occitan dans tous ses aspects (langue, littérature, civilisation, musique, ethnologie...). Tous les trois ans, un grand congrès international, en France ou à l'étranger, permet de rassembler l'ensemble de ces chercheurs.

Question 5

Quel est l'état actuel de la recherche et de l'enseignement de l'occitan à Montpellier? Et quelles actions et pistes à partir du centre de Montpellier ont été pensées pour l'avenir, en termes de dialogue avec la société sur l'occitan?

L'université de Montpellier est une des premières en France à avoir fait de l'occitan un objet de recherche dès la fin du XIXe siècle. L'équipe de chercheurs occitanistes publie d'ailleurs aujourd'hui encore la Revue des Langues Romanes²³, fondée en 1870 par Camille Chabanneau qui a occupé la première chaire de langues romanes de Montpellier. C'est Charles Camproux, un des membres fondateurs de la Societat d'Estudis Occitans, préfiguration de l'actuel Institut d'Estudis Occitans, qui a créé le premier cours de langue et littérature occitane en 1942. Disciple de Camproux, Robert Lafont (1923-2009) est ensuite devenu une personnalité incontournable dans l'histoire de la recherche en occitan dans la mesure où ses travaux ont couvert des domaines aussi divers que la littérature médiévale, la littérature contemporaine, la linguistique. Professeur d'occitan à l'université Paul-Valéry, il a orienté une grande partie de ses recherches vers la sociolinguistique. Après avoir fondé le Groupe d'Étude et de Recherche sur la Diglossie Franco-Occitane (GERDFO), l'équipe de chercheurs montpelliérains a fondé en 1977 la revue de sociolinguistique *Lengas*²⁴.

Le centre de recherches en domaine occitan (RedOc) de l'université de Montpellier est le seul en France et à l'étranger dont les travaux se concentrent presque exclusivement sur la langue occitane, même si cette équipe travaille par ailleurs sur d'autres axes de recherches transversaux au sein de l'équipe universitaire ReSO²⁵ (Recherches sur le Suds et les Oriens) dont elle est une composante.

L'équipe RedOc/ReSO est directement liée à un cursus de formation universitaire complet en occitan. Elle se compose de quatre enseignants-chercheurs titulaires et de deux enseignants issus du secondaire (PRAG) pour la partie enseignement, mais leurs compétences dans des domaines aussi variés que la médiévistique, la linguistique (dialectologie, phonologie), l'histoire

22 <http://www.aieo.org>.

23 <https://journals.openedition.org/rlr/>.

24 <https://journals.openedition.org/lengas/>.

25 <https://reso.www.univ-montp3.fr>.

(médiévale, moderne et contemporaine) et la littérature (moderne et contemporaine) permettent de couvrir l'ensemble du champ spatio-temporel de la langue et de la littérature d'oc.

L'ensemble des chercheurs de l'équipe RedOc est impliqué dans un parcours universitaire d'enseignement de l'occitan qui forme des spécialistes de la langue (enseignants du primaire et du secondaire, chercheurs, archivistes...). Il s'agit de la seule équipe qui soit étroitement liée à un dispositif pratique de formation en occitan en France. Tous les chercheurs de l'équipe interviennent donc à tous les niveaux de la formation. Cette formation propose une licence d'occitan, deux masters dont un master recherche et un master MEEF qui prépare les étudiants aux métiers du professorat (cursus bilingue, CAPES et agrégation d'occitan). Les formations en occitan de l'université sont également entièrement disponibles en France et à l'étranger grâce au service d'enseignement à distance de l'université. Le développement de l'enseignement de l'occitan est fortement dépendant des dispositifs institutionnels mis en place au niveau national et ce développement reste pour l'heure relativement limité dans la mesure où le nombre de postes ouvert aux différents concours de l'enseignement reste très faible (quatre postes par an pour le CAPES, un poste pour l'agrégation). En revanche, dans la mesure où l'image de l'occitan profite d'une nette amélioration dans la société française depuis quelques années, la formation en occitan ouvre aujourd'hui de nouvelles perspectives professionnelles pour les étudiants dans des domaines aussi variés que les collectivités territoriales, le tourisme, le journalisme, les bibliothèques et dans divers métiers liés à la culture.

Question 6

Pourriez-vous aborder la dynamique des mouvements de revitalisation des langues et la contribution de la sociolinguistique et des domaines de connaissance connexes ? Quel rôle les perspectives idéologiques jouent-elles dans cette dynamique ? Quand les mouvements pour la revitalisation ou le renforcement des langues minoritaires dépassent-ils les idéaux et les projets politiques pour atteindre la réalité sociale des pratiques communicatives? Dans quelle mesure les communautés de langue occitane servent-elles de normes de référence pour les espaces officiels (d'éducation, de documentation, de communication, de développement technologique)?

Les mouvements renaissantistes du XIXe siècle, au premier rang desquels le Félibrige étaient tournés majoritairement vers l'écriture littéraire. Frédéric Mistral, figure de proue de ce mouvement a d'ailleurs obtenu le Prix Nobel de littérature en 1904. À ce premier mouvement de revitalisation a succédé l'Institut d'estudis occitans (IEO), après la Seconde Guerre mondiale, qui en promouvant une norme orthographique nouvelle, basée sur les principes de la graphie médiévale, et en soutenant la création littéraire, a œuvré pour faire de l'occitan non

plus seulement un objet d'étude et de création, mais également un outil de communication et de pensée moderne. Cet institut a orienté peu à peu son action vers la reconquête sociale de la langue. L'école qui avait contribué à la substitution linguistique a représenté alors l'espoir d'une reconquête possible de la langue. Un des acquis institutionnels majeurs du XX^e siècle pour l'occitan comme pour toutes les autres langues de France sera donc la promulgation de la Loi Deixonne (1951) qui a autorisé pour la première fois l'enseignement de l'occitan dans les écoles. Ainsi l'occitan a pu peu à peu se développer à tous les niveaux du système éducatif, de la maternelle à l'université.

La mise en circulation du concept de diglossie dans la seconde moitié du XX^e siècle a suscité un intérêt particulier du côté des chercheurs spécialistes des langues régionales, en particulier du côté catalan et occitan. La situation de ces langues dominées face à des langues nationales dominantes a fait émerger une nouvelle approche du concept de diglossie. Dans le cadre d'une sociolinguistique occitano-catalane dite « périphérique », des « chercheurs natifs » ont considéré que si rien n'était fait concrètement pour défendre la langue, la situation diglossique ne pouvait conduire à terme qu'à l'assimilation des langues dominées, autrement dit à la disparition du catalan et de l'occitan au profit respectivement de l'espagnol et du français. Ce constat les a conduits à mettre en œuvre tout un travail de normalisation et de revalorisation de ces langues devant conduire à revitaliser leurs usages dans les registres, notamment formels, dont elles avaient été progressivement privées. À une vision statique de la diglossie qui supposait une répartition stable et équilibrée des fonctions linguistiques entre deux formes de langue, Robert Lafont (1997) a opposé une vision dynamique et polémique rendant compte de la position de subordination de la langue occitane qui, comme en témoigne l'histoire, s'est vue progressivement privée de ses registres d'expression les plus prestigieux (littéraires ou officiels) au profit du français. Au XX^e siècle c'est la survie même de la langue occitane qui était en jeu avec la disparition progressive de ses usages oraux et l'arrêt de la transmission familiale. Pour Lafont, si rien n'était entrepris pour défendre et reconquérir les fonctions linguistiques perdues par l'occitan, la langue était condamnée à disparaître. Refusant la neutralité, l'approche proposée par Robert Lafont (1984) assume le développement d'une sociolinguistique engagée, socialement et politiquement, et une implication sociale pour la défense de la langue occitane puisque les processus diglossiques sont perçus comme indissociables de la mise en place d'une idéologie nationale dominante.

Au-delà de la question sociale de la langue, de son avenir et de sa normalisation, l'approche scientifique devait donc permettre avant tout de donner un statut et une dignité à une langue occitane toujours vivante et à un objet d'étude occulté pendant trop longtemps derrière la dénomination de « patois ». Dans la continuité de cette approche, l'équipe RedOc s'attache donc aujourd'hui en particulier à rétablir la continuité entre les usages littéraires médiévaux, qui seuls avaient droit de cité à l'université, et les usages contemporains (écrits ou oraux) qui n'y

avaient jusqu'à une époque récente que peu ou pas de place. Alors que l'on a longtemps opposé « provençal ancien » (nom donné autrefois à l'occitan médiéval) et « patois » (terme péjoratif que les locuteurs-mêmes de l'occitan ont fini par intégrer), les travaux sociolinguistiques du XXe siècle ont permis d'envisager la langue occitane dans une perspective diachronique ininterrompue (de son origine à son statut actuel de langue régionale). Les conditions d'existence de la langue occitane en situation de diglossie, de sa communication écrite et orale, produisant des effets sur la langue qu'il est possible de décrire, c'est à cette tâche que se consacrent en grande partie les chercheurs de l'équipe RedOc.

Question 7

Quelles actions et interactions en termes de sociolinguistique militante ou d'un activisme sociolinguistique prospèrent ou ont le potentiel de prospérer dans la prise de conscience et le soutien institutionnel de la relation entre la science linguistique de l'occitan et la société ? Dans quelle mesure les mouvements et documents de l'Unesco ont-ils contribué à l'attention portée aux langues minoritaires ou minorisées, notamment pour la langue occitane ?

Au cours du XXe siècle, de nombreuses actions ou manifestations ont été organisées par des associations, notamment culturelles, pour faire reconnaître l'occitan comme langue et lui redonner une place sociale. L'engagement des chercheurs occitanistes pour la défense de l'occitan a sans doute contribué aux progrès institutionnels que connaît l'occitan aujourd'hui dans la société française, même si ces progrès restent encore relativement faibles. Dans tous les cas, le début d'institutionnalisation de l'occitan en France et le soutien de plus en plus qu'il obtient au niveau des administrations régionales montre que l'image négative de l'occitan, considéré autrefois comme un « patois », s'est nettement améliorée, à tel point que l'appellation « occitan » est aussi utilisée aujourd'hui à des fins commerciales ou touristiques (cf. Garabato ; Boyer, 2020). De plus en plus de villes et villages choisissent aujourd'hui de restituer la toponymie en occitan (noms de villes et de villages, noms de rues...), signe que la langue occitane est perçue comme un élément important du patrimoine.

Au-delà des nombreuses créations artistiques en occitan (théâtre, musique, chansons, littérature) qui trouvent encore un écho auprès des occitanophones anciens, mais aussi auprès des jeunes générations de néolocuteurs, les divers mouvements de défense de l'occitan se sont structurés en organismes qui aujourd'hui contribuent à la revitalisation de l'occitan. L'enseignement s'est peu à peu institutionnalisé et structuré à tous les niveaux de la formation. Tous les enseignants de l'enseignement public des 33 départements français concernés se retrouvent au sein de la Fédération des enseignants de langue et culture occitane (FELCO)²⁶ et produisent des ressources pédagogiques de qualité pour l'enseignement.

26 <http://www.felco-creo.org>.

Ces progrès de l'image de l'occitan et son institutionnalisation de plus en plus importante ne doivent pas occulter le fait que la question de la survie de cette langue n'est pas réglée. À l'heure actuelle le nombre de locuteurs reste insuffisant pour assurer la survie de la langue, notamment en l'absence de transmission familiale. Aujourd'hui les jeunes générations ont de moins en moins contact avec des locuteurs de l'occitan. Les derniers locuteurs natifs disparaissent peu à peu et l'enjeu pour les prochaines années est de parvenir à former suffisamment de néoluteurs pour que la langue ne disparaisse pas et continue de se transmettre. C'est la raison pour laquelle le développement de l'enseignement de l'occitan par immersion ou dans des parcours bilingues représente à l'heure actuelle un enjeu essentiel pour l'avenir de l'occitan.

Références

BEC, P. *La langue occitane*. Que sais-je? n° 1059. Paris : Presses Universitaires de France, 1995 [1963, 1ère édition].

Bec, P. Per una dinamica novèla de la lenga de referéncia : Dialectalitat debasa e diasistèma occitan, 39-61, *Annales de l'institut d'études occitanes*, n°6, T2, Nîmes, IEO, 1972.

FERGUSON, C. A. *Diglossia*. *Word* 15, p. 325-340, 1959.

FISHMAN, J. Bilingualism with and without Diglossia; Diglossia with and without Bilingualism. *Journal of Social Issues*, 23, 29-38, 1967.

GARABATO, C. A. ; BOYER, H. *Le marché et la langue occitane au vingt-et-unième siècle : microactes glottopolitiques contre substitution*. Limoges : Lambert Lucas, 2020.

LAFONT, R. *Quarante ans de sociolinguistique à la périphérie*. Paris : l'Harmattan, 1997.

LAFONT, R. Pour retrouver la diglossie, *Lengas* n°15, 1984.



**LUGAR SOCIAL E VARIAÇÃO
DIASSISTÊMICA DO OCCITANO**

Entrevista com Hervé Lieutard

**PLACE SOCIALE ET VARIATION
DIASYSTÉMIQUE DE L'OCCITAN**

Interview avec Hervé Lieutard

Hervé Lieutard

Marcia dos Santos Machado Vieira¹

Vanessa Meireles

RESUMO

Hervé Lieutard é professor titular de linguística occitana na Universidade Paul-Valéry – Montpellier 3 desde 2001. Sua pesquisa concentra-se na análise da variação dialetal do occitano a partir de uma perspectiva fonológica, mas também sobre o lugar social do occitano ao longo de sua história, desde a época medieval até os dias de hoje, estudada através da vasta documentação escrita disponível neste idioma. Nesta entrevista, ele faz um breve panorama das dificuldades da pesquisa em occitano ligadas ao declínio social do occitano, mas também mostra como os processos de revitalização contribuíram para mudar a imagem da língua e para desenvolver novamente sua transmissão e socialização. ²

PALAVRAS-CHAVE: Occitano; Língua regional; Diglossia; Sociolinguística; Fonologia.

1 *Marcia dos Santos Machado Vieira* (marcia@letras.ufrj.br) é professora-pesquisadora da Faculdade de Letras da Universidade Federal do Rio de Janeiro, vinculada ao Programa de Pós-Graduação *stricto sensu* em Letras Vernáculas (Língua Portuguesa) da mesma universidade, e, entre outras responsabilidades, coordena o Projeto franco-brasileiro VariaR – Variação em Línguas Românicas – em parceria com *Vanessa Meireles* (vanessa.meireles@univ-montp3.fr), professora de Língua Portuguesa, Cultura e Literatura do Brasil na Universidade Paul Valéry - Montpellier 3 (França) desde 2017, membro da equipe de pesquisa *Recherches sur les Suds et les Orient* (ReSO).

2 Tradução para o português da entrevista original em francês, por Vanessa Meireles, e revisão final por Marcia dos Santos Machado Vieira.

Questão 1

Professor Hervé Lieutard, poderia falar-nos brevemente de seu perfil de interesses de investigação e articulação de pesquisa(s)?

A língua occitana tem sido o foco da minha pesquisa por mais de vinte anos. Meu método de estudo desta língua regional da França combina tanto uma abordagem fonológica da variação dialetal específica desta língua como uma abordagem sociolinguística que permite uma melhor compreensão dos efeitos do contato linguístico, em particular entre o occitano e o francês, sobre as mudanças que esta língua sofreu ao longo de sua história e sobre a evolução de seus usos e seu lugar na sociedade francesa.

O occitano, apesar de seu desenvolvimento particularmente importante na Idade Média e seu prestígio literário, nunca se tornou a língua oficial de um Estado. Na ausência de uma forma padrão oficial, é, portanto, uma língua pluridialetoal que se caracteriza por uma variação linguística significativa na área em que é falada (por volta de um terço da França, ou seja, 33³ departamentos – divisão administrativa do país –, mas também está presente em quatorze vales do Piemonte italiano e no Vale de Aran, na Espanha). Além da variação dialetal que pode ser observada nesta língua, as estruturas comuns aos vários dialetos permitem definir um diassistema occitânico no qual a intercompreensão estabelece os limites externos entre as seis principais variedades que compõem esta língua (*Languedocien, Provençal, Gascon, Limousin, Auvergnat e Vivaro-Alpin*). A abordagem fonológica que desenvolvo em minha pesquisa é alimentada por dados da dialectologia de campo. Há vários anos, venho tentando mostrar que o uso de teorias fonológicas sobre a sílaba em constituintes organizados não só nos permite dar conta da estrutura interna desta língua, mas também de sua variação dialetal, especialmente no nível de terminações silábicas. Esta abordagem nos permite ir além da tradicional abordagem geolinguística da variação baseada em isoglossas. A tipologia silábica que eu defini para o occitano serve como uma unidade de medida da variação dialetal, ao mesmo tempo em que torna possível dar conta do funcionamento supradialetoal que reflete a unidade subjacente desta linguagem além da evolução dessemelhante de seus dialetos no tempo e no espaço. A diversidade dialetal desta língua ainda representa um vasto campo de investigação para fonólogos. Nos últimos anos, também ampliei meu campo de pesquisa, interessando-me pela tipologia silábica das línguas românicas, com especial atenção ao catalão, uma língua que gradualmente se separou do occitano a partir da Idade Média por razões históricas, mas que ainda é totalmente compreensível para os falantes de occitano.

Outro aspecto de minha pesquisa diz respeito ao surgimento de padrões linguísticos escritos no período medieval, para o qual o occitano sem dúvida representa um dos modelos incomparáveis no domínio românico. O occitano tendo sido pioneiro no estabelecimento de práticas de escrita padronizadas na Idade Média, tanto no registro lírico dos Trovadores,

3 39 se contarmos os departamentos nos quais ela está apenas parcialmente presente.

como também para os usos pragmáticos da linguagem (tratados retóricos, gramáticas, textos jurídicos, notariais, administrativos ou científicos). Eu também estou trabalhando na edição digital de textos antigos em occitano em formatos padronizados (XML-TEI) que agora nos permitem prever a interoperabilidade entre os diversos documentos editados neste formato com o objetivo de propor um processamento parcialmente automatizado dos dados obtidos (análise morfosintática e lexical, assim como análise gráfica). Estes trabalhos permitem avaliar as diversas evoluções que caracterizam o occitano desde o período medieval até os dias atuais e, em particular, avaliar e medir a continuidade do uso léxico entre a Idade Média e o período contemporâneo.

A abordagem sociolinguística que caracteriza minha pesquisa está ligada ao fato de que é impossível fazer pesquisa sobre o occitano sem também se interessar pelo lugar social desta língua e pelos efeitos do contato linguístico sobre seu uso. A revista *Lengas*⁴, que dirijo há vários anos, é uma revista de sociolinguística que dá um lugar importante a esta questão do contato linguístico, não só para o occitano, mas também para todas as outras línguas em situação de diglossia. Estes trabalhos se inscrevem na esteira dos modelos teóricos desenvolvidos pela escola de sociolinguística de Montpellier em torno de Robert Lafont, que, em colaboração com sociolinguistas catalães, está na origem de uma análise original da situação das línguas dominadas ou em contato, em particular através de uma releitura crítica do conceito de diglossia proposto por Ferguson (1959) e Fishman (1967). Esta perspectiva sociolinguística tem guiado minha pesquisa sobre a questão da escrita occitana desde a Idade Média até os dias de hoje, seja para o estudo da literatura occitana ou para usos escritos não literários. Nos últimos anos, também dediquei uma série de estudos aos usos gráficos do occitano em outros períodos, na medida em que eles são de alguma forma um espelho do lugar do occitano na sociedade francesa, particularmente no período do século XV ao século XIX, que se caracteriza por um declínio social desta língua.

Questão 2

Qual é o estatuto linguístico e social do occitano: língua materna, segunda língua? É língua de quantos falantes, de que comunidade(s), de que tipos de textos orais e escritos (de documentos e/ou veículos oficiais, por exemplo)?

O occitano sofre há muito tempo com a falta de visibilidade na sociedade francesa devido aos preconceitos que envolvem sua prática e uma política linguística que, desde a Revolução Francesa em particular, tem tentado erradicar o “*patois*” (num sentido pejorativo) e generalizar o uso do francês nas práticas orais. Na França, as representações dos idiomas foram e ainda são em grande parte moldadas pelo modelo da língua francesa, que foi imposto institucionalmente e que levou a uma confusão entre a hierarquia administrativa oficial dos idiomas (língua nacional

4 <https://journals.openedition.org/lengas/>.

e língua regional) e uma hierarquia de qualidade desses idiomas. Este não é o espaço para definir o longo processo de minoração do occitano na sociedade francesa, mas vale a pena lembrar apenas que a construção do francês oficial, particularmente desde a Revolução, se desenvolveu através de um discurso de negação da alteridade linguística que consistiu em particular em desconstruir ideologicamente todo o espaço cultural e linguístico do occitano. Embora o occitano tenha sido uma língua medieval de prestígio, seu uso administrativo escrito começou a declinar no século XVI com a oficialização do francês escrito (Portaria de Villers-Cotterêts em 1539, sob o reinado de François I). Seguindo as políticas linguísticas levadas a cabo pelo Estado francês para promover o francês oral, particularmente desde a Revolução Francesa, a transmissão familiar diminuiu gradualmente, mas foi sobretudo com a introdução da escolarização obrigatória no final do século XIX (Leis Jules Ferry de 1880-1881) que o conhecimento do francês por todos se tornou obrigatório. Os movimentos renascentistas e literários occitanos que surgiram no século XIX (*Félibrige*, Occitanismo...) não foram capazes de reverter esta tendência, a ponto de a transmissão familiar do occitano ter parado em meados do século XX.

Entre as várias línguas da França, agora é mais fácil identificar o bretão, o corso, o basco e até mesmo o catalão porque essas línguas também levam o nome de uma região, enquanto o occitano, que nunca se constituiu como uma entidade política distinta, leva apenas o nome de sua língua (occitano ou *langue d'oc* vem da maneira como se diz “sim” nessa língua, ou seja, “òc”).

Hoje é difícil estimar com exatidão o número de falantes do occitano na medida em que a construção da identidade francesa em torno da ideia de um povo indivisível não permite que sejam feitas perguntas nas estatísticas realizadas na França em nível nacional que poderiam levar a pensar que existem identidades ou grupos étnicos distintos. No início do século XX, o linguista Jules Ronjat estimou que existiam 10 milhões de falantes do occitano. A pesquisa realizada pelo *Ofici public de la lenga occitana* (OPLO) em 2020 em duas das quatro grandes regiões de língua occitana (Occitânia e Nouvelle-Aquitaine) indica que o número de locutores que declaram falar occitano sem dificuldade corresponde a aproximadamente 7% da população desses territórios⁵. Como a transmissão de família para família parou gradualmente nas décadas seguintes à introdução da escolarização obrigatória, os últimos falantes que herdaram o occitano como língua materna são muito idosos. O occitano agora só é transmitido como segunda língua, principalmente no contexto escolar.

Em todos os casos, a língua occitana tem uma grande riqueza de documentação cobrindo todos os períodos, desde a Idade Média até os dias de hoje. Embora os cancioneiros dos Trovadores tenham sido objeto de numerosos estudos de pesquisa na França e no exterior,

5 <https://www.ofici-occitan.eu/fr/restitution-des-resultats-de-lenquete-sociolinguistique/>.

muito poucos estudos foram dedicados à escrita pragmática, administrativa ou científica em occitano durante este período, embora os documentos sejam muito abundantes. O trabalho de digitalização e publicação digital realizado na França e na Europa permite cada vez mais o livre acesso a vários documentos desde a Idade Média (Dicionário informatizado do occitano medieval⁶, textos literários⁷, textos de occitano pragmático⁸, etc.) até os dias de hoje (<https://occitanica.eu/enciclopedia>). Apesar do status diaglósico da língua, que limitou severamente seu desenvolvimento, o uso oral e literário do occitano nunca foi interrompido até hoje. Um movimento literário renascentista occitano começou no século XVI, depois novamente no século XIX com a fundação do *Félibrige*⁹ em torno do poeta Frédéric Mistral, que também escreveu o grande dicionário enciclopédico e etnográfico que ainda hoje é uma referência, *Lou Tresor dóu Felibrige*¹⁰. Ainda hoje existe uma produção literária significativa na Occitânia e várias editoras que publicam obras em occitano (prosa¹¹, poesia¹²...). Há também várias estações de rádio occitanas¹³, alguns jornais, especialmente em formato eletrônico¹⁴, e alguns programas em occitano em canais de televisão regionais¹⁵.

De um ponto de vista lexicográfico, os primeiros dicionários do Occitano começaram a aparecer no século XVIII. Os trabalhos de coleta linguística começaram no século XIX com a criação dos primeiros atlas e os primeiros estudos dialectológicos. Pesquisas ainda estão sendo realizadas hoje entre os últimos falantes nativos. As várias campanhas de pesquisa resultaram no mapeamento de todo o território de língua occitana. Os dados de campo coletados nos vários atlas produzidos entre os séculos XIX e XXI estão agora reunidos no *Thesoc*¹⁶, um banco de dados computadorizado desenvolvido por uma equipe de pesquisadores em Nice, que contém todos os dados dialetais sobre a Occitânia. A documentação oral em occitano é importante, mas estes dados estão dispersos em vários centros¹⁷. Entretanto, algumas delas foram recentemente processadas para a análise da microvariação sintática nos idiomas da França¹⁸. Em todo caso, o mais importante centro de documentação na Occitânia é Lo CIRDOC – *Institut occitan de*

6 <http://www.dom-en-ligne.de>.

7 Por exemplo, o repertório digitalizado da literatura trovadoresca, Rialto (<http://www.rialto.unina.it/>).

8 Por exemplo, a edição recente dos registros do Consulado Medieval de Montpellier (<http://thalamus.huma-num.fr>).

9 <http://www.felibrige.org/>.

10 <https://gallica.bnf.fr/ark:/12148/bpt6k74854.image>.

11 Por exemplo, IEO-editions, https://ideco-dif.com/ieo_editions.

12 Por exemplo, as Edições Jorn, <http://www.editions-jorn.com/catalogue.htm>.

13 Por exemplo, Ràdio Lengadòc, <https://radiolengadoc.com>.

14 Por exemplo, Jornalet, <https://www.jornalet.com>.

15 Por exemplo Viure al País, <https://france3-regions.francetvinfo.fr/occitanie/emissions/viure-al-pais-0>.

16 <http://thesaurus.unice.fr/>.

17 Por exemplo Talvera, <http://cordae-talvera-documentation.kentika.fr>.

18 <http://symila.univ-tlse2.fr>.

*cultura*¹⁹. Trata-se de um centro associado à Biblioteca Nacional da França (BNF) que tem como objetivo reunir todos os recursos e conhecimentos do campo occitano com um objetivo enciclopédico, acessível em particular através de seu portal multimídia *Occitanica*²⁰.

Questão 3

Como o governo francês e a sociedade francesa em geral veem o idioma occitano, seu uso e seu ensino? Que políticas nacionais de governo, políticas públicas, existem para promoção e valorização dessa língua na França e além da França?

O status oficial do occitano hoje é múltiplo, pois é falado em três Estados diferentes. Na Espanha, no Vale de Aran, o occitano é co-oficial com o catalão e o castelhano e como tal é utilizado oficialmente na administração, de acordo com o Estatuto de Autonomia da Catalunha, mas esta oficialidade do occitano na Catalunha diz respeito apenas a uma pequena parte do número total de falantes de occitano, já que há menos de 10.000 habitantes no Vale de Aran. A proporção de ensino na Occitânia nesta região aumenta para 60% até o início do ensino médio, mas depois diminui. Na Itália, nos vales do Piemonte, o occitano tem o status de língua legalmente protegida, mas o italiano continua sendo a única língua oficial e o occitano não é ensinado nas escolas. Na França, o status linguístico do occitano é simplesmente o de uma língua regional. O artigo 75-1 da Constituição francesa, votado em 2008, estipula que “as línguas regionais fazem parte do patrimônio da França”, mas na medida em que apenas o Francês é a língua oficial da República (artigo nº 2 da Constituição: “a língua da República é o francês”), seu uso é limitado aos círculos privados (famílias, amigos etc.), embora seu ensino tenha sido autorizado desde 1951 desde o jardim de infância até a universidade (lei Deixonne). Desde aquela época, no entanto, o ensino do occitano se desenvolveu amplamente no ensino público (aulas de idioma, turmas bilíngues), mas também no ensino privado (ensino imersivo).

As posições oficiais do governo francês não favorecem atualmente o desenvolvimento do occitano em grande medida. A Carta Europeia das Línguas Regionais e Minoritárias, elaborada pelo Conselho da Europa para assegurar a proteção das línguas históricas regionais ou minoritárias da Europa, foi assinada pelo governo francês em 1999, mas, embora a assinatura da Carta tenha sido baseada em uma seleção de medidas moderadas, o texto foi finalmente rejeitado pelo Conselho Constitucional francês, porque, de acordo com suas declarações, “A Carta Europeia das Línguas Regionais ou Minoritárias contém cláusulas contrárias à Constituição” e atinge “os princípios da indivisibilidade da República, da igualdade perante a lei e da singularidade do povo francês” (Decisão nº 99-412 DC de 15 de junho de 1999). Ao conferir direitos específicos às comunidades linguísticas, a Carta Europeia, de acordo com o Conselho Constitucional, seria, portanto, contrária ao primeiro artigo da Constituição Francesa, que proclama a indivisibilidade

19 <https://www.oc-cultura.eu>.

20 <https://occitanica.eu/>.

da República, mas seria também, segundo o Conselho, difícil de conciliar com o artigo nº 2, que proclama que “a língua da República é o francês”.

Apesar desses bloqueios, que impedem o reconhecimento oficial genuíno do occitano, alguns avanços institucionais devem ser destacados. Desde 2016, um acordo entre o Estado francês e as regiões da Occitânia e Nouvelle-Aquitaine permitiu o estabelecimento do *Ofici public de la lenga occitana* (OPLO)²¹, uma organização que ajuda a promover o occitano na sociedade, apoiando a formação de professores e estudantes e o desenvolvimento do occitano nas escolas. Ela também subsidia projetos que contribuem para a transmissão e socialização da língua occitana. Infelizmente, esta organização cobre apenas duas das quatro principais regiões nas quais o occitano é uma língua histórica. Desde 2011, *Lo Congrès*²² também oferece ferramentas linguísticas on-line (dicionários, conjugadores, bases de dados terminológicas, digitalização de dicionários históricos, aplicações de síntese e reconhecimento de voz etc.). Atualmente, sou vice-presidente do Conselho Linguístico deste órgão de regulamentação linguística apoiado pelo Ministério francês da Cultura. Um projeto para um grande dicionário histórico da língua occitana está sendo desenvolvido atualmente dentro da estrutura deste órgão.

Questão 4

Como o professor perspectiva a inserção dessa língua no conjunto de línguas românicas? Que atenção em termos de práticas de pesquisa, ensino e difusão (inclusive, via caminho de ciência aberta) tem recebido em contextos universitários e científicos voltados a línguas românicas?

Além do fato de que o occitano é uma língua ameaçada de extinção segundo a UNESCO, ainda hoje faz parte das “*understudied languages*”, em outras palavras, línguas que são subestudadas e para as quais ainda há muita pesquisa a ser feita a fim de melhor compreendê-las. O reconhecimento científico de um espaço linguístico occitano tem sido lento e, como resultado, o occitano só muito recentemente se tornou um objeto de estudo científico por direito próprio. Nas últimas décadas do século XIX, quando se desenvolveu um novo interesse pela literatura ocidental medieval dos trovadores e começaram os primeiros trabalhos científicos de linguística sobre o occitano, vários representantes da linguística francesa oficial, em primeiro lugar Gaston Paris e Paul Meyer, ambos professores do *Collège de France*, desenvolveram amplamente um discurso de negação e rejeição da própria noção do espaço linguístico do occitano e suas variedades dialetais, a fim de melhor apoiar o projeto de unificação nacional em torno do francês e de sua capital, considerando que as variedades occitanas não eram nada mais que variedades distantes e “deformadas” de francês, que eram pejorativamente chamadas de

21 <https://www.ofici-occitan.eu/fr/accueil/>.

22 <https://locongres.org/oc/>.

“*patois*”. Por esta forma de desprezo, eles contribuíram para invalidar e despojar os linguistas do sul de seu objeto de trabalho e contribuíram para reforçar todos os preconceitos transmitidos naquela época sobre as línguas da França, privando os falantes de occitano da possibilidade de acesso a uma forma de “consciência linguística” occitana. Em apoio a estas teorias ideológicas de unificação nacional, foi desenvolvida a noção de Galo-România, que é de fato uma aplicação inadequada dos conceitos histórico-geográficos da linguística e que mostra acima de tudo o peso da ideologia na construção de fronteiras linguísticas arbitrárias. A reconstrução e legitimação de um espaço linguístico do occitano exigiu, antes de mais nada, um longo e provavelmente ainda inacabado processo de desconstrução dos preconceitos e representações, tanto endógenos quanto exógenos, aos quais a língua foi submetida.

No entanto, o princípio da intercompreensão occitana sempre permaneceu na base de todos os estudos linguísticos e sociolinguísticos dedicados ao occitano nos séculos XX e XXI. Foi o linguista Pierre Bec quem adaptou a noção de diassistema ao occitano, propondo uma nova estruturação do occitano que foi além dos limites dialetais tradicionais definidos antes dele. Na última edição de seu trabalho de referência sobre a língua occitana (BEC, 1995, 6), Pierre Bec estendeu a noção de diassistema a todo o espaço que ele definiu como occitano-romance, incluindo o occitano e o catalão neste conjunto distinto de línguas galo-românicas. Na perspectiva de Pierre Bec, o diassistema tornou possível ver nos dialetos occitano e catalão em sincronia não como emergências de novas formas linguísticas, mas sim evoluções díspares no tempo e no espaço de um único protosistema linguístico original (BEC, 1972, 46).

Seja como for, o occitano ocupa um lugar central no conjunto das línguas românicas e é sem dúvida uma das pontes mais interessantes para as outras línguas românicas meridionais (italiano, catalão, castelhano, até mesmo português...). Infelizmente, na ausência de uma forma oficialmente reconhecida e de uma política oficial do Estado francês para sua promoção, esta língua permanece em grande parte desconhecida, ainda hoje em seu próprio território e, em particular, pelos linguistas romanistas. É sobretudo a literatura medieval que tem despertado o interesse dos pesquisadores em nível internacional, mesmo se hoje em dia cada vez mais pesquisas estão sendo feitas sobre literatura moderna e contemporânea, civilização e linguística. Não há muitos pesquisadores trabalhando sobre o occitano na França, mas felizmente existe uma rede internacional de pesquisadores agrupados na Associação Internacional de Estudos Occitânicos (AIEO)²³, da qual sou secretário geral, que reúne cerca de 300 pesquisadores e acadêmicos de todo o mundo interessados em occitano em todos os seus aspectos (língua, literatura, civilização, música, etnologia etc.) A cada três anos, um grande congresso internacional, na França ou no exterior, reúne todos esses pesquisadores.

23 <http://www.aieo.org>

Questão 5

De que maneira se dá a pesquisa e o ensino atual sobre o occitano em Montpellier? E que ações e caminhos a partir do centro de Montpellier têm sido pensados/estão no horizonte, em termos de interlocução com a sociedade sobre o occitano?

A Universidade de Montpellier foi uma das primeiras na França a tornar o occitano um tema de pesquisa no final do século XIX. Além disso, a equipe de pesquisadores occitanistas publica a *Revue des Langues Romanes* (Revista de Línguas Românicas)²⁴, fundada em 1870 por Camille Chabanneau, que ocupou a primeira cátedra de línguas românicas em Montpellier. Foi Charles Camproux, um dos membros fundadores da *Societat d'Estudis Occitans*, o precursor do atual *Institut d'Estudis Occitans*, que criou o primeiro curso de língua e literatura occitana em 1942. Um discípulo de Camproux, Robert Lafont (1923-2009) tornou-se posteriormente uma figura chave na história da pesquisa occitana, pois seu trabalho cobriu campos tão diversos como a literatura medieval, a literatura contemporânea e a linguística. Como professor de occitano na Universidade de Paul-Valéry, ele concentrou grande parte de suas pesquisas em sociolinguística. Após a fundação do *Groupe d'Étude et de Recherche sur la Diglossie Franco-Occitane* (Grupo de Estudo e Pesquisa sobre a diglossia Franco-Occitânica – GERDFO), a equipe de pesquisadores de Montpellier fundou a revista sociolinguística *Lengas*²⁵ em 1977.

O *Centre de recherches en domaine occitan* (Centro de pesquisa da área occitânica – RedOc) da Universidade de Montpellier é o único na França e no exterior cujo trabalho se concentra quase exclusivamente na língua occitana, embora esta equipe também trabalhe em outros eixos transversais de pesquisa dentro da equipe universitária ReSO (*Recherches sur le Suds et les Oriens* – Pesquisas sobre o Sul(s) e o Oriente(s)) da qual faz parte.

A equipe RedOc/ReSO está diretamente ligada a um programa completo de treinamento universitário na Occitânia. É composto por quatro professores-pesquisadores titulares e dois professores do ensino médio (*PRAG*) para a parte de ensino, mas suas habilidades em campos tão variados como estudos medievais, linguística (dialectologia, fonologia), história (medieval, moderna e contemporânea) e literatura (moderna e contemporânea) permitem-lhes cobrir todo o campo espaço-temporal da língua d'oc e sua literatura.

Todos os pesquisadores da equipe RedOc estão envolvidos em um curso universitário sobre o ensino do occitano, que forma especialistas da língua (professores do ensino fundamental e médio, pesquisadores, arquivistas etc.). É a única equipe que está intimamente ligada a um programa de treinamento prático na Occitânia, na França. Todos os pesquisadores da equipe estão, portanto, envolvidos em todos os níveis de treinamento. Este programa oferece uma

24 <https://journals.openedition.org/rlr/>

25 <https://journals.openedition.org/lengas/>.

licenciatura em occitano, dois mestrados, incluindo um mestrado em pesquisa e um mestrado em ensino que prepara os estudantes para cargos de ensino (currículo bilíngue, concursos franceses para professor em nível de ensino médio *CAPES* e *Agrégation* em occitano). Os diplomas em occitano da universidade também estão totalmente acessíveis na França e no exterior através do ensino à distância da universidade. O desenvolvimento do ensino do occitano está fortemente dependente dos arranjos institucionais colocados em prática em nível nacional e este desenvolvimento permanece por enquanto relativamente limitado na medida em que o número de postos abertos aos diversos concursos de ensino permanece muito baixo (quatro postos por ano para o concurso *CAPES*, um posto para o concurso *Agrégation*). Por outro lado, na medida em que a imagem do occitano melhorou consideravelmente na sociedade francesa nos últimos anos, a formação em occitano abre agora novas perspectivas profissionais para estudantes em áreas tão variadas como autoridades/coletividades locais, turismo, jornalismo, bibliotecas e em várias profissões relacionadas à cultura.

Questão 6

Poderia abordar a dinâmica de movimentos de revitalização de línguas e a contribuição da Sociolinguística e de áreas do conhecimento correlacionadas? Que papel têm as perspectivas ideológicas nessa dinâmica? Quando movimentos em prol da revitalização ou fortalecimento de línguas minorizadas ultrapassam ideais e projetos políticos e chegam à realidade social das práticas comunicativas? Em que medida comunidades de fala do occitano servem de norma de referência a espaços oficiais (de educação, documentação, comunicação, desenvolvimento tecnológico)?

Os movimentos renascentistas do século XIX, em cuja vanguarda estava Félibrige, preocupavam-se principalmente com a escrita literária. Frédéric Mistral, a figura principal deste movimento, recebeu o Prêmio Nobel de Literatura em 1904. Este primeiro movimento de revitalização foi sucedido após a Segunda Guerra Mundial pelo *Institut d'estudis occitans* (IEO), que, ao promover um novo padrão ortográfico baseado nos princípios da ortografia medieval e ao apoiar a criação literária, trabalhou para tornar o occitano não apenas um objeto de estudo e criação, mas também uma ferramenta de comunicação e pensamento moderno. Este instituto orientou gradualmente sua ação para a recuperação social da língua. A escola, que contribuiu para a substituição linguística, representou a esperança de uma possível reconquista do idioma. Uma das maiores realizações institucionais do século XX para o occitano, como para todas as outras línguas da França, foi a promulgação da Lei Deixonne (1951) que autorizou o ensino do occitano nas escolas pela primeira vez. Assim, o occitano foi capaz de se desenvolver pouco a pouco em todos os níveis do sistema educacional, desde o jardim de infância até a universidade.

A circulação do conceito de diglossia na segunda metade do século XX despertou particular interesse entre os pesquisadores das línguas regionais, especialmente o catalão e o occitano. A situação destas línguas dominadas diante das línguas nacionais dominantes deu origem a uma nova abordagem do conceito de diglossia. No âmbito da chamada sociolinguística “periférica” occitano-catalã, os “pesquisadores nativos” consideraram que se nada fosse feito concretamente para defender a língua, a situação de diglossia só poderia levar a longo prazo à assimilação das línguas dominadas, ou seja, ao desaparecimento do catalão e do occitano em benefício do espanhol e do francês, respectivamente. Esta observação levou-os a implementar todo um processo de normalização e revalorização destas línguas, o que deveria levar à revitalização de seu uso nos registros, particularmente os formais, dos quais haviam sido progressivamente privadas. Em contraponto a uma visão estática da diglossia, que assumiu uma distribuição estável e equilibrada das funções linguísticas entre duas formas de linguagem, Robert Lafont (LAFONT, 1997) apresentou uma visão dinâmica e polêmica que levava em conta a posição subordinada da língua occitana, que, como a história mostra, foi progressivamente privada de seus registros de expressão mais prestigiados (literários ou oficiais) em favor do francês. No século XX, a própria sobrevivência da língua occitana estava em jogo com o desaparecimento gradual de seus usos orais e o fim da transmissão familiar. Para Lafont, se nada fosse feito para defender e recuperar as funções linguísticas perdidas pelo occitano, o idioma estava condenado a desaparecer. Recusando a neutralidade, a abordagem proposta por Robert Lafont assume o desenvolvimento de uma sociolinguística engajada social e politicamente e um envolvimento social para a defesa da língua occitana, uma vez que os processos de diglossia são percebidos como inseparáveis do estabelecimento de uma ideologia nacional dominante.

Além da questão social da língua, seu futuro e sua normalização, a abordagem científica deveria permitir acima de tudo dar um status e uma dignidade a uma língua occitana ainda viva e a um objeto de estudo que está escondido há muito tempo por trás da denominação de “*patois*”. Na continuidade desta abordagem, a equipe RedOc está hoje particularmente preocupada em restabelecer a continuidade entre os usos literários medievais, que por si só tinham um direito de lugar na universidade, e os usos contemporâneos (escritos ou orais), que até recentemente tinham pouco ou nenhum lugar. Enquanto durante muito tempo houve uma oposição entre “antigo provençal” (nome dado ao occitano medieval antigamente) e “*patois*” (um termo pejorativo que os próprios falantes do occitano acabaram integrando), os trabalhos de sociolinguística do século XX tornaram possível considerar a língua occitana numa perspectiva diacrônica ininterrupta (desde sua origem até seu status atual como língua regional). As condições de existência da língua occitana em uma situação de diglossia, de sua comunicação escrita e oral, produzindo efeitos na língua que é possível descrever, é a esta tarefa que os pesquisadores da equipe RedOc dedicam amplamente seus trabalhos.

Questão 7

Que ações e interações em termos de uma Sociolinguística ativista ou de um ativismo sociolinguístico prosperam ou têm o potencial de prosperar para sensibilização e apoio institucional à relação entre ciência linguística do occitano e sociedade? Em que medida movimentos e documentos da Unesco têm contribuído para atenção a línguas minoritárias ou minorizadas, especialmente para o occitano?

Durante o século XX, numerosas ações ou eventos foram organizados por associações, particularmente culturais, para que o occitano fosse reconhecido como uma língua e para dar-lhe um lugar social. O engajamento dos pesquisadores occitanistas com a defesa do occitano contribuiu sem dúvida para os progressos institucionais que o occitano está fazendo na sociedade francesa de hoje, mesmo que estes progressos ainda sejam relativamente fracos. De qualquer forma, o início da institucionalização do occitano na França e o crescente apoio que está recebendo ao nível das administrações regionais mostra que a imagem negativa do occitano, antes considerado como um *patois*, melhorou claramente, tanto que o nome ‘occitano’ agora também é usado para fins comerciais ou turísticos (cf. GARABATO; BOYER, 2020). Cada vez mais cidades e povoações estão optando por restaurar sua toponímia em occitano (nomes de cidades e aldeias, nomes de ruas etc.), um sinal de que a língua occitana é percebida como uma parte importante de sua herança.

Além das numerosas criações artísticas em occitano (teatro, música, canções, literatura) que ainda ressoam junto aos falantes de occitano mais velhos, mas também com as gerações mais jovens de neo-locutores, os vários movimentos de defesa do occitano foram estruturados em organizações que hoje contribuem para a revitalização do occitano. O ensino tornou-se gradualmente institucionalizado e estruturado em todos os níveis de formação. Todos os professores das escolas públicas dos 33 departamentos franceses envolvidos são membros da *Fédération des enseignants de langue et culture occitane* (Federação de professores de língua e cultura occitana – FELCO)²⁶ e produzem recursos didáticos de qualidade para o ensino.

Este progresso na imagem do occitano e sua crescente institucionalização não devem obscurecer o fato de que a questão da sobrevivência da língua não foi resolvida. Atualmente, o número de falantes permanece insuficiente para garantir a sobrevivência do idioma, particularmente na ausência de transmissão familiar. Hoje, as gerações mais jovens têm cada vez menos contato com falantes de occitano. Os últimos falantes nativos estão desaparecendo pouco a pouco e o desafio para os próximos anos é treinar neo-locutores suficientes para que o idioma não desapareça e continue a ser transmitido. É por isso que o desenvolvimento do ensino de occitano através de imersão ou formações bilíngues é atualmente uma questão essencial para o futuro do occitano.

26 <http://www.felco-creo.org>

Referências

BEC, P. *La langue occitane*. Que sais-je? n° 1059. Paris: Presses Universitaires de France, 1995 [1963, 1ère édition].

BEC, P. Per una dinamica novèla de la lenga de referéncia: Dialectalitat debasa e diasistèma occitan, 39-61, *Annales de l'institut d'études occitanes*, n°6, T2, Nîmes, IEO, 1972.

FERGUSON, C. A. *Diglossia*. *Word* 15, p. 325-340, 1959.

FISHMAN, J. Bilingualism with and without Diglossia; Diglossia with and without Bilingualism. *Journal of Social Issues*, 23, 29-38, 1967.

GARABATO, C. A.; BOYER, H. *Le marché et la langue occitane au vingt-et-unième siècle: microactes glottopolitiques contre substitution*. Limoges: Lambert Lucas, 2020.

LAFONT, R. *Quarante ans de sociolinguistique à la périphérie*. PARIS: L'HARMATTAN, 1997.

LAFONT, R. Pour retrouver la diglossie, *Lengas* n°15, 1984.



**AQUISIÇÃO/ENSINO DO PORTUGUÊS EUROPEU
POR/A FALANTES ARABÓFONOS EM MARROCOS:
ESPECIFICIDADES E ELEMENTOS DE RESPOSTA¹**

**ACQUISITION/TEACHING OF EUROPEAN PORTUGUESE
BY/TO ARABIC SPEAKERS IN MORROCO: SPECIFICITIES AND
RESPONSE ELEMENTS**

Maria Antónia Mota²

Habiba Naciri³

RESUMO

A primeira Licenciatura em Estudos Portugueses no mundo árabe teve início no ano académico de 2009-10, em Marrocos, na Universidade Mohammed V de Rabat (UM5R), ao abrigo de um Protocolo de Cooperação entre a UM5R, a Universidade de Lisboa (ULisboa) e o Camões-Instituto da Cooperação e da Língua (Camões I.P.). A Licenciatura tornou-se autónoma, no final do ano letivo 2020-21, mantendo-se, contudo, a cooperação entre a UM5R e a Faculdade de Letras da Universidade de Lisboa (FLUL), no âmbito das áreas de interesse mútuo, no qual se inclui a língua e a cultura portuguesas/de expressão portuguesa, nas suas variedades. No presente artigo, enquadra-se a Licenciatura do ponto de vista académico e do ponto de vista sociolinguístico – chamando-se a atenção para o multilinguismo social que caracteriza Marrocos e se espelha na população estudantil, com grande diversidade nas combinatórias de línguas previamente adquiridas e iniciando o estudo do português em idade adulta. Com vista a adequar o ensino do português, na sua variedade nacional europeia e enquanto língua adicional, maioritariamente, a uma população universitária com características linguísticas peculiares, está em curso o Projeto MAPEAR (Construção de materiais de apoio ao ensino e à aquisição tardia de PLE por falantes arabófonos), que se descreve, apresentando-se alguns dados linguísticos ilustrativos dos domínios linguísticos que requerem especial atenção, retirados do *corpus* de produções escritas e orais já obtidas junto dos estudantes da Licenciatura, acompanhados de uma reflexão sobre o que esses dados nos dizem sobre as interlínguas e nos recomendam, em termos de ação.

PALAVRAS-CHAVE: Aquisição/ensino do português europeu; Língua adicional; *Corpus* de produções escritas e orais; Materiais para docentes e estudantes; O projeto MAPEAR.

1 Este artigo decorre da apresentação feita pelas autoras no I Seminário Internacional VariaR - *O ensino do português no mundo*, realizado online em 28/10/2021, com os naturais acrescentos de conteúdo. Faz-se notar que H. Naciri teve um papel particularmente relevante na informação disponibilizada sobre a situação linguística em Marrocos.

2 *Maria Antónia Mota* é professora associada aposentada da Faculdade de Letras da Universidade de Lisboa, investigadora do Grupo Dialectologia & Diacronia, mantonia.mota@letras.ulisboa.pt.

3 *Habiba Naciri* é professora no departamento de estudos portugueses da Faculdade de Letras e Ciências Humanas, Universidade Mohammed V-Rabat, investigadora e membro permanente do laboratório Línguas, tradução, culturas e comunicação, habiba.naciri@gmail.com.

ABSTRACT

The first University Degree in Portuguese Studies in the Arab world began in the academic year of 2009-10, in Morocco, at the Mohammed V University of Rabat (UM5R), under a Cooperation Protocol between the UM5R, the University of Lisbon (ULisboa) and Camões-Institute for Cooperation and Language (Camões IP). Portuguese Studies became autonomous at the end of the 2020-21 academic year. Auspiciously, the cooperation remains between UM5R and the Faculty of Arts of the University of Lisbon (FLUL), throughout a specific protocol on mutual scientific interests, which includes the Portuguese language and culture/the language and culture of Portuguese expression, in its varieties. In this article, the Graduate Degree is framed from an academic point of view, and from a sociolinguistic point of view - drawing attention to the social multilingualism that characterizes Morocco and is mirrored in the student population, with great diversity in the combinations of previously acquired languages and beginning the study of Portuguese in adulthood. In order to adequate the teaching of Portuguese, in its European national variety (mostly, as additional language), to a university population with peculiar linguistic characteristics, the MAPEAR Project (Construction of materials to support teaching and the late acquisition of PLE by Arabic-speakers) will be presented, as well as some linguistic data, taken from the corpus of written and oral productions obtained from the students of the Graduate Degree, accompanied by a reflection on what the data tell us about interlanguages, and about the appropriate actions.

KEYWORDS: European Portuguese acquisition/teaching; Additional language; Corpus of written and oral productions; Materials for teachers and students; The MAPEAR project.

Introdução

No ano académico de 2009-10, entrou em funcionamento a primeira *Licenciatura em Estudos Portugueses* no chamado Mundo Árabe, sendo, por uma década, a única nesse grande espaço geográfico e cultural.⁴ A sua abertura foi proposta à Universidade de Lisboa pela Presidência da Universidade Mohammed V de Rabat (UM5R), enquadrando-se no Protocolo de Cooperação já existente entre as duas Universidades. Tendo essa proposta sido muito bem recebida, iniciaram-se os trabalhos de elaboração de um plano de estudos para seis semestres⁵ e dos programas do conjunto das unidades de crédito que o iriam integrar, definindo-se os objetivos de cada uma das cadeiras e indicando-se a bibliografia de base requerida. Esse trabalho foi levado a cabo na ULisboa⁶ e posteriormente discutido e aprovado pelos responsáveis da Faculdade de Letras e Ciências Humanas (FLSH) da UM5R.

4 Uma Licenciatura em Estudos Portugueses e Lusófonos foi aberta, em 2019, na Faculdade Al Alsun da Universidade Ain Shams, no Cairo, ao abrigo de um protocolo com a Universidade do Porto e o Camões- Instituto da Cooperação e da Língua.

5 A estrutura desta Licenciatura segue os pressupostos do Processo de Bolonha, sendo o 1º Ciclo (Licenciatura) de seis semestres/três anos letivos, o qual é geralmente completado pelos estudantes com o 2º Ciclo (Mestrado). A Declaração de Bolonha, assinada em 1999, hoje com 48 signatários, está na origem de um sistema que visa harmonizar o ensino na Europa, criando o Espaço Europeu de Ensino Superior, entretanto alargado pela adoção do sistema por países não europeus.

6 Em 2008-09, a Professora Inês Duarte, então Vice-Reitora, liderou o processo, pela Reitoria da ULisboa; Maria Antónia Mota foi nomeada representante da ULisboa e coordenadora científica da Licenciatura pela mesma Universidade, tendo participado no referido processo desde o início. O nosso principal interlocutor, à época, foi o Professor Jamal E. Elhani, desde há vários anos Decano da FLSH e nosso interlocutor direto.

O referido plano de estudos tinha como objetivo oferecer uma formação universitária especializada no domínio do português. Dando particular relevo ao domínio linguístico, visava igualmente uma formação cultural o mais abrangente possível, tendo em conta não só a abertura a uma língua e realidade novas, por parte dos futuros estudantes, mas também o âmbito específico em que a formação se inseria. A capacitação dos licenciados para aceder a pós-graduações de interesse para o seu futuro profissional ou diretamente a diferentes atividades profissionais foi igualmente considerada. Assim, os 6 semestres foram estruturados em torno de três eixos: (i) formação em língua, linguística e tradução/retroversão; (ii) desenvolvimento das aquisições pela prática, sob forma de oficinas de compreensão e de expressão orais e escritas (nomeadamente, com recurso a documentos áudio e vídeo autênticos e à leitura de periódicos *online*, preferencialmente sobre temas associados ao terceiro eixo, a seguir); (iii) formação em literaturas escritas em português (de Portugal, de África, do Brasil), em cultura e sociedade (manifestações culturais e da vida social em Portugal, com elementos também associados a espaços não europeus de língua oficial portuguesa), em história da língua portuguesa e das suas atuais variedades, em factos históricos marcantes da sociedade portuguesa (focando, nomeadamente, as relações históricas entre Portugal e Marrocos) e em arte portuguesa. Uma relação entre aspetos históricos e manifestações artísticas foi também tida em conta, na perspetiva da articulação de saberes; por exemplo, no período dos sécs. XVI-XVIII, plasmanse na arte portuguesa elementos da história dos Descobrimentos e da época colonial, e em séculos posteriores, do colonialismo florescente e do Estado Novo português, seguido da implantação da democracia. Partindo do pressuposto de que um paralelo cronológico entre estas áreas do conhecimento beneficia a aquisição de uma visão mais harmoniosa e atrativa das diferentes épocas, as disciplinas de cada semestre referem-se todas a um mesmo período, mais ou menos lato. Os 6 semestres visam, assim, uma progressão na complexidade dos conteúdos e dos seus suportes, a articulação entre as diferentes áreas do conhecimento e o diálogo entre o reforço dos conhecimentos anteriormente adquiridos e a introdução de novos elementos. Entretanto, algumas reestruturações foram introduzidas ao plano inicial, em função de diretivas emanadas do Ministério da Educação marroquino, mas mantiveram-se os princípios norteadores desta formação.

Por proposta da ULisboa, o Camões-Instituto da Cooperação e da Língua (Camões I.P.) integrou um Protocolo tripartido que enquadrou os doze primeiros anos de vida da Licenciatura, assinado para o efeito e em vigor durante o período de 2009-10 a 2020-21. Dispondo de uma dupla coordenação, em íntima articulação – científica, pela representante da ULisboa, e pedagógica, pelo(a) responsável pela FLSH da UM5R –, a Licenciatura tornou-se, recentemente, autónoma, cumprindo-se assim o objetivo delineado desde o início da cooperação. O Departamento de Português foi, entretanto, legalmente criado, com a constituição de um corpo docente marroquino permanente, cujo doutoramento, na área do português, se efetivou já no quadro desta cooperação.

A Universidade de Lisboa desempenhou um papel científico fundamental: para além de se responsabilizar pela definição das bases de funcionamento da Licenciatura, como acima referido,

autorizou a cooperação de doutorados da ULisboa, reconhecidos especialistas nas diferentes áreas, sob forma de docência graciosa, em períodos concentrados, de todas as unidades de crédito para as quais não existia massa crítica na UM5R. Muitos dos professores da ULisboa que participaram neste projeto foram e são orientadores de teses de Mestrado e de Doutoramento de licenciados em Rabat. O nosso reconhecimento para com esses colegas é imenso, sendo muito aprazível constatar que todos eles consideraram esta cooperação muito estimulante. Assim, no início, deslocaram-se em missão de ensino 15 docentes da ULisboa (Faculdade de Letras e Faculdade de Belas-Artes), tendo o seu número progressivamente diminuído, à medida que se foram doutorando colegas marroquinos e que vários licenciados progrediram nos estudos: se em curso de Doutoramento, os estudantes são autorizados a assegurar algumas horas semanais de trabalho docente, o que se reveste da maior importância – com efeito, essa é uma maneira de se irem formando especialistas em diferentes áreas de interesse para a Licenciatura e para futuros programas de pós-graduação, na UM5R⁷.

A ULisboa disponibilizou-se, ainda, para assinar convénios com a UM5R para elaboração de teses de Doutoramento em regime de cotutela e a FLUL assinou com a FLSH de Rabat um convénio *Erasmus+*, o primeiro naquela faculdade marroquina e que já deu frutos, com a frequência de um semestre por 6 licenciandos de Rabat na FLUL. Com o apoio financeiro da Associação para o Desenvolvimento da Faculdade de Letras de Lisboa, organizaram-se já dois cursos de atualização, na FLUL, destinados aos docentes marroquinos da Licenciatura.

O apoio do Camões I.P., por seu lado, permitiu a presença continuada (e ainda atual) de um leitor/docente português na Licenciatura, o que favorece um contacto regular com um falante nativo, por parte de estudantes e docentes marroquinos. Esse apoio concretizou-se, ainda, durante os referidos doze anos, em relevantes apoios financeiros de várias ordens, com destaque para a atribuição de um número importante de bolsas de investigação para os estudantes que seguiram para Mestrado e Doutoramento, na ULisboa, e para o financiamento das despesas associadas às missões de ensino dos docentes da ULisboa e da coordenadora representante da mesma Universidade.

A conjugação de esforços permitiu que, hoje, exista uma biblioteca do Departamento de português, com um número expressivo de obras, devidamente catalogadas, de teor científico, de literatura, de arte, etc., adquiridas com subsídios do Camões I.P. e através de ofertas (por exemplo, da Fundação C. Gulbenkian e de docentes da ULisboa). Os estudantes frequentam esse espaço, sendo convidados a fazer empréstimos de livros para lerem durante as férias – de referir que foram reunidos livros infantojuvenis portugueses, com características que permitam que os estudantes das fases iniciais os possam ler com facilidade, como ajuda para a sua progressão de conhecimentos linguísticos. A existência de um Centro de Língua Portuguesa, do Camões I.P., bastante bem equipado, é um fator muito importante para o apoio do Departamento e da sua missão.

7 No início, a Licenciatura contou também com um segundo professor português, de língua, residente em Casablanca.

Na fase atual, mantém-se, em boa hora, a cooperação protocolada entre a FLSH-UM5R e a FLUL, no âmbito das áreas de interesse mútuo, nelas se destacando a língua e a cultura portuguesas/de expressão portuguesa. Nesse quadro, diferentes atividades estão previstas, também na área da História (relações diplomáticas, património português construído em Marrocos, etc.), destacando-se, para efeitos do presente artigo, o lançamento do projeto MAPEAR, que adiante se apresentará.

De destacar, ainda, nesta visão panorâmica da cooperação em causa, a importância que sempre teve, e continua a ter, o inestimável apoio do Decanato da Faculdade de Letras de Rabat e, a outro nível, o da Embaixada de Portugal em Rabat, que, sempre que solicitados, intervieram no sentido da facilitação e do bom rumo das iniciativas necessárias, apoiando e estimulando os responsáveis pela coordenação da Licenciatura⁸.

Enquadramento da Licenciatura em Estudos Portugueses

A par do interesse científico e social da abertura de uma Licenciatura com o perfil acima descrito no Mundo Árabe e da colaboração interinstitucional que a enquadrou, o que valeria por si só, faz-se notar o alcance de que se reveste esta colaboração, como elemento constitutivo da política de cooperação Norte-Sul (Europa-Magrebe). Na Declaração de Barcelona adotada na Conferência Euro-mediterrânica, de 1995, na qual participou Marrocos, são defendidos aspetos da maior relevância para a paz e o desenvolvimento, sendo que os signatários se declararam "convictos de que o objetivo geral [é o] de tornar a bacia mediterrânica numa zona de diálogo, de intercâmbio e de cooperação que garanta a paz, a estabilidade e a prosperidade"⁹.

Citamos, ainda, o seguinte, pela relevância do papel da sociedade civil, na qual se incluem as universidades: "a sociedade civil pode contribuir de forma essencial para o processo de desenvolvimento da parceria Euro-Mediterrânica", criando ou fomentando "os necessários instrumentos de uma cooperação (...) para incentivar as trocas entre os agentes do desenvolvimento: responsáveis da sociedade civil e política, mundo cultural e religioso, universidades, investigadores, meios de comunicação social, associações, sindicatos e empresas públicas e privadas"¹⁰ e visando "estabelecer [esta] parceria (...) através (...) da atribuição de uma maior importância

⁸ Ao longo do tempo, os Embaixadores de Portugal em Rabat, para além do relevante papel diplomático na defesa da língua e cultura portuguesas em Marrocos, sempre organizaram, assessorados pelos seus Conselheiros Culturais, no Centro Cultural da Embaixada, cursos de português abertos à sociedade, muito procurados e ministrados por sucessivos leitores/docentes Camões e/ou por licenciados em Estudos Portugueses com aptidões adequadas à função.

Uma palavra é também devida às representações diplomáticas do Brasil, de Angola e da Guiné-Bissau em Rabat, que sempre aceitaram os nossos convites para participar nas comemorações anuais do Dia da Língua Portuguesa por nós organizadas, assim como ao Presidente da Universidade de El-Jadida, que acolheu uma dessas comemorações. Não esquecemos, também, os bolseiros dos PALOP em Rabat, que confraternizaram por algumas vezes com os nossos estudantes, nesses dias de festa, e que assim ajudaram a fortalecer o conhecimento sobre as variedades linguísticas e as manifestações culturais de diferentes espaços onde se fala português.

⁹ https://www.europarl.europa.eu/summits/mad4_pt.htm

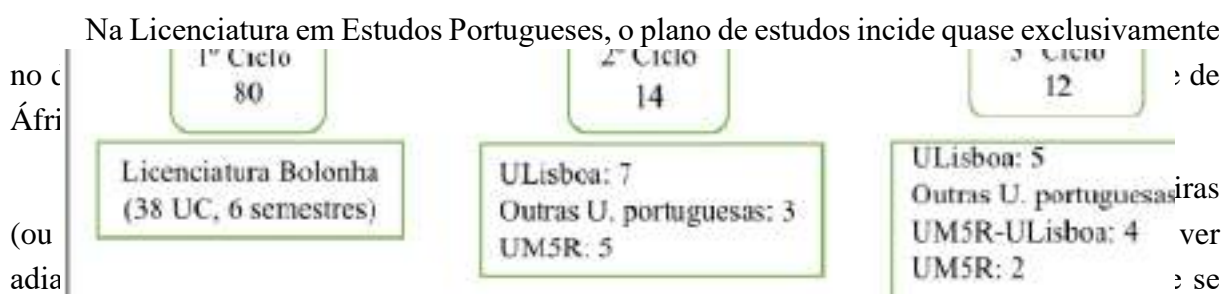
¹⁰ https://www.europarl.europa.eu/summits/mad1_pt.htm#citiz

à dimensão social, cultural e humana, constituindo estes três eixos as três vertentes da parceria Euro-Mediterrânica".

Se as relações entre Portugal e Marrocos datam de há séculos, a sua configuração é bem distinta, nas últimas décadas, verificando-se um incremento das relações comerciais, do turismo, da agricultura e pescas, da indústria (construção civil, energia, geologia e minas), do meio ambiente, entre outras, assim como de cooperação nas áreas da segurança no Mediterrâneo e Atlântico. Este desenvolvimento tem consequências diretas na empregabilidade dos nossos estudantes, como abaixo se destacará.

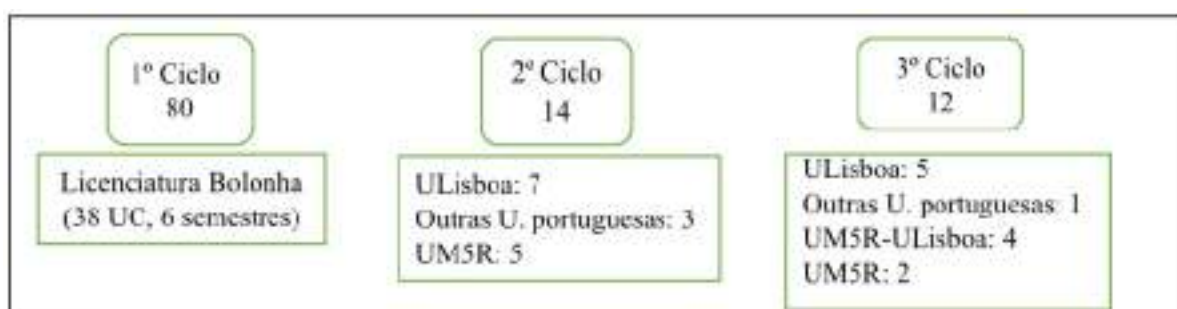
Principais resultados da formação em Estudos Portugueses

Na UM5R, são atualmente ensinadas muitas línguas estrangeiras, existindo várias licenciaturas congêneres da Licenciatura em Estudos Portugueses, entre as quais três centradas em línguas românicas, as de Espanhol, Francês e Italiano.



mostra na Fig. 1, um número importante dos cerca de 80 licenciados em Estudos Portugueses continuou a sua formação com a obtenção de um Mestrado, tendo vários deles seguido para Doutoramento; noutros casos, e por os estudantes (cinco) já terem obtido um Mestrado noutras áreas, em Marrocos, a passagem a Doutoramento na área do português fez-se diretamente¹². Na mesma Figura, indicam-se as universidades em que essa formação decorreu ou decorre, ainda no presente, sendo que se trata maioritariamente de universidades portuguesas, com largo destaque para ULisboa:

Figura 1 – Formação dos estudantes nos três Ciclos de estudos



11 A língua árabe e a francesa são também objeto de ensino formal, com dois semestres cada, como nas restantes licenciaturas em línguas.

12 Nos números indicados, incluem-se 1 dissertação de Mestrado e 4 teses de Doutoramento em fase final ou mesmo de defesa pública.

As áreas de pós-graduação científica são, em termos gerais, as da língua e linguística, das ciências da linguagem, da tradução, das literaturas de expressão portuguesa, da história, da arte. Os cinco estudantes que obtiveram o seu Mestrado em Rabat, fizeram-no na Escola Superior de Educação de Rabat, na área da didática das línguas estrangeiras.

Parte das dissertações e teses beneficia da orientação apenas de professores de instituições universitárias portuguesas e outra parte, de orientações conjuntas ULisboa-UM5R. Deve sublinhar-se que uma parte substancial dos estudantes obteve bolsas de investigações do Camões I. P., o que permitiu a mobilidade dos estudantes e a sua permanência por longos períodos em Lisboa, tendo, assim, proximidade com os seus orientadores portugueses, acesso a bibliotecas e, igualmente muito importante, total imersão linguística e social; outros, porém, prosseguiram a sua formação a expensas próprias, juntando a verba necessária com o esforço do seu trabalho, prévio ou concomitante, o que mostra bem o enorme empenho dos estudantes em se especializarem.

Os resultados das pós-graduações têm sido excelentes, com notas muito elevadas, sendo que duas doutorandas foram convidadas a integrar centros de investigação portugueses e um outro foi chamado a lecionar árabe, num curso livre do Centro de Línguas da FLUL (CLi). Um outro índice muito positivo da aquisição de uma formação em português é o da taxa de empregabilidade de licenciados e pós-graduados, que chega a $\geq 90\%$, em empresas nacionais e internacionais, embaixadas de países de língua oficial portuguesa, Polícia, ensino, tradução, turismo, jornalismo, *calling centres*, etc.

Em Marrocos, o ensino é totalmente gratuito, pelo que vários estudantes obtêm duas licenciaturas, ou até dois mestrados. Assim sendo, o público universitário apresenta um perfil muito variado, ao entrar na Licenciatura em Estudos portugueses, não só no que respeita à formação universitária prévia como à faixa etária, juntando-se estudantes saídos do ensino secundário, com cerca de 17 anos, e pessoas com formação superior noutras áreas, com idade (muito) superior. Em comum têm o facto de iniciarem a aquisição do PLE apenas na universidade (o português não faz parte do elenco das LE oferecidas em níveis de ensino anteriores), o que se tenta compensar com um mês de formação intensiva em língua, antes do início do 1º semestre da Licenciatura. Se tal exige um enorme esforço por parte de estudantes (aprendentes tardios) e docentes, não impede que aqueles estudantes que ultrapassam com sucesso a barreira do 1º ano¹³ se tornem maioritariamente bons alunos. Em secção posterior, voltaremos a estas questões.

Contextualização sociolinguística do ensino/aquisição do português em Marrocos: o multilinguismo marroquino

13 No 1º ano, há tradicionalmente muitos inscritos nesta Licenciatura, nos Serviços Académicos. Mas há, igualmente, uma prática de mudança de curso, mesmo antes do início das aulas, e uma grande taxa de absentismo, pelo que a diferença entre inscritos nos Serviços no 1º ano da Licenciatura e o número efetivo de presenças e de frequência regular é muito grande.

Marrocos é um país norte-africano de duas expressões nacionais oficiais: o árabe e o *amazigh*, conhecido na literatura internacional sob o nome de berbere. Mas o *amazigh* foi declarado língua oficial apenas em 2011, na nova Constituição marroquina.

Apesar de o espaço deste artigo não permitir elencar todo o processo de estabelecimento do árabe nas terras marroquinas dos *imazighen* (berberes), consideramos relevante fornecer alguma informação histórica, de modo a melhor se compreender a atual situação linguística do país e da sua população.

Assim, o árabe clássico penetrou nas terras marroquinas à medida que foi levada a cabo a arabização do território, mas, sobretudo, a islamização. De facto, a religião fez com que este idioma fosse facilmente acolhido e substituísse o *amazigh* em muitos domínios da vida da população autóctone. Pela grande carga afetiva que tem, e mesmo não sendo inicialmente língua materna, foi desde então usado não só no domínio da liturgia e da aprendizagem do Corão como na expressão literária, continuando a ser descrito na maioria das gramáticas. Aceitando a religião muçulmana, os *imazighen* deixaram espaço à língua do Corão para que esta obtivesse o estatuto de língua oficial. Não obstante, esta população conservou a sua identidade linguística e cultural através das gerações, por via oral, mesmo não podendo levar esta língua à oficialidade até recentemente. Como variedades maternas dos descendentes dos *imazighen*, com uma expressão bastante grande no país, contam-se

- (i) o *tachelhit*, variedade com maior número de falantes, muito utilizada no sul de Marrocos, particularmente na cidade de Marraquexe, passando por Agadir, Taroudante, Ouarzazate, Essaouira e os arredores destas;
- (ii) o *tamazight*, usado no centro do país, nas regiões rurais em torno de cidades como Khenifra, Fez, Ifran, Khemisat, etc.;
- (iii) o *tarifit*, variedade do noroeste de Marrocos, na região do Rif, em cidades como Tetuan, Nador, Husaima, entre outras.

Por seu lado, no domínio da língua árabe, existem variedades diferenciadas não só linguisticamente, por traços específicos, mas quanto aos domínios de uso. Estas variedades vão desde a variedade alta (árabe clássico standard) à variedade baixa (árabe marroquino), desenvolvido neste espaço geográfico: de facto, ninguém usa o árabe clássico, no dia a dia, mas sim o árabe marroquino, língua materna de uma vasta parte da população, conhecido sob o nome do *darija*. O *darija* é, assim, a língua materna dos descendentes dos árabes marroquinos. Mas, num movimento de diferenciação social das pessoas instruídas, sobretudo em situações do uso formal, surgiu, entretanto, um árabe marroquino que denominamos “culto/standard” no Quadro 1, abaixo, chamado ‘mediano’, na terminologia de Youssi (1983), ou ‘árabe marroquino moderno’, em Ennaji (2001), aos quais voltaremos abaixo.

Quadro 1 – Panorama linguístico de Marrocos

L1		L apenas de aprendizagem formal	L veicular urbana
<i>amazigh</i> ↓ <i>tachelhit</i> <i>tamazight</i> <i>tarifit</i>	<i>darija</i> ou árabe marroquino comum árabe marroquino moderno / culto / standard	- árabe clássico ⇒ uso litúrgico (mesquita) - árabe (clássico) ⇒ modalidade formal standard	francês espanhol

O árabe (clássico) standard é a língua do Estado e das suas instituições e de comunicação com outras regiões árabes. Para Youssi (1983), o árabe, em Marrocos, passou de uma situação de diglossia (árabe clássico – árabe marroquino) a uma situação de triglossia, com a inclusão da nova variedade. Já Ennaji (2001), passados cerca de vinte anos sobre a publicação anterior, considera existir quadriglossia, em que cada variedade escrita e oral se duplica, dando origem a outro registo linguístico, incluindo o desenvolvimento do “árabe moderno” – ensinado na escola e que levanta muitos problemas de aquisição às crianças que só conhecem o *darija* e/ou uma variedade do *amazigh*, sem acesso à língua clássica no seu meio.¹⁴ Baseando-nos nos trabalhos acima, apresentamos a seguinte representação:

Quadro 2 – Situação do árabe no contexto marroquino

	Língua árabe			
	árabe clássico (escrito)		árabe marroquino (oral)	
Diglossia	árabe clássico (escrito)		árabe marroquino (oral)	
Triglossia (Youssi, 1983)	árabe clássico (/literário)		árabe mediano	árabe marroquino
Quadriglossia (Ennaji, 2001)	árabe clássico	árabe moderno	árabe mediano	árabe marroquino

Temos, assim, um *continuum* social de línguas/variedades, com diglossia, sendo que, como é sabido, o valor da língua/variedade depende das circunstâncias da sua utilização, do seu peso e prestígio social. Bennis (2011: 3, 4) refere o seguinte: “[houve] uma mudança imprevisível que se reflete no estabelecimento de uma variedade urbana cuja principal função é estabelecer o *continuum* formal e nocional que faltava, um *continuum* entre o árabe marroquino, o *amazigh* marroquino e o árabe padronizado. (...) O retorno do árabe marroquino aos meios de comunicação é, sem dúvida, o fenómeno mais marcante no mercado linguístico marroquino (...) [como se verifica, por exemplo] nos painéis colocados na esfera pública (...). Todos estes elementos constituem espaços de promoção e enriquecimento da língua materna a que se refere o termo ‘marroquino’” (t.n.).

14 Trata-se de uma variedade simplificada do árabe clássico, resultante do movimento Nahda, ou Renascimento, entre finais do séc. XVIII e 1950, “era romântica, mesclando cultura e política. Corresponde a vários movimentos de emancipação: do pensamento e da linguagem que o exprime, (...) da “nação árabe” no contexto otomano ou colonial, da libertação também das mulheres” (<https://www.monde-diplomatique.fr/mav/106/DUPONT/17685>; t.n.). Citando Youssi (1983: 78), até à data “nenhum governo ou instituição árabe propôs a adoção de variedades nacionais como línguas oficiais e de escolarização”.

Podemos dizer que, na realidade, o árabe clássico sofreu mudanças exigidas pelas circunstâncias sociopolíticas e pelas grandes inovações tecnológicas que o mundo árabe conheceu nos séculos anteriores. No século XX, depois da independência dos países árabes, estes adotaram uma política de arabização do ensino. Através do processo da tradução, palavras novas penetraram o léxico árabe e estruturalmente assistiu-se a processos de mudança sistemática¹⁵, resultando daí a versão simplificada do árabe moderno ou padrão/standard, que passou a usar-se essencialmente nos meios audiovisuais e na comunicação social.

O último censo realizado em Marrocos data de 2014, não havendo, assim, dados atualizados; mas, nesse Censo, o uso era de 89,96% para o *darija* e de 25,74% para o ramo *amazigh*, em média; na província de Rabat, onde se localiza a UM5R, o uso do *darija* é aí estimado em cerca de 98% e o do *amazigh*, em cerca de 12%. Naturalmente, nestas percentagens incluem-se os falantes que falam duas ou mais destas línguas¹⁶. Assim, na sociedade marroquina, podem encontrar-se dentro do mesmo espaço geográfico falantes monolíngues, bilingues ou multilingues.

Uma pergunta impõe-se: no mapa linguístico marroquino, que língua oficial é realmente considerada nas comunicações oficiais? Como é que uma língua recém-codificada, como é o caso do *amazigh*, se torna uma língua de comunicação para uma massa que a herdou pelo berço sem se instruir nela? E que variedade do *amazigh* é considerada efetivamente na comunicação oficial? No começo do século XXI, Marrocos abriu uma nova página de conciliação com o *amazigh*, ao proceder à sua oficialização, como já referido, a inserção no ensino, até agora em escolas modelo, sabendo-se que se ensina uma versão unida das suas três variedades (o *amazigh* standard), ou simplesmente, uma nova língua. O próprio nome da ortografia *amazigh*, *tifnagh*, determina esta finalidade: tifi (achado)+nagh (nosso), ou seja, ‘o nosso achado’, quer dizer, a forma encontrada para salvar o património *amazigh*. A nosso ver, esta reconciliação com a língua *amazigh* é de aplaudir; porém, nota-se precipitação em algumas decisões assumidas. Pensou-se em conservar a identidade *amazigh* e esqueceu-se que a língua *amazigh* nos chegou de geração em geração apenas oralmente, sendo que muitos falantes dessa língua não conseguem compreender a sua versão escrita unificada, o *tifnaghe*, já que não conhecem as convenções ortográficas correspondentes.¹⁷

15 Ennaji (2001: 55), comparando o árabe clássico e o árabe moderno, afirma que “o mesmo tipo de frases é usado tanto no árabe clássico quanto no árabe padrão (árabe moderno) com uma diferença: os morfemas na posição final de cada palavra estão ausentes no árabe padrão” (t.n.). A nível da ordem das palavras na frase, é frequente o uso da estrutura S(ujeito) V(erbo) O(bjeto), diferentemente da ordem VSO do árabe clássico; na sintaxe, pelo contacto com o francês, muitas formas desta língua penetraram no árabe moderno. Ennaji (2001: 56) justifica esta inovação dando o exemplo do sintagma agente; para este investigador, “O sintagma agente no pretérito min qibali, ‘alaa jadi, min laduni’ (em nome de), que anteriormente não existia em árabe, é amplamente usado em árabe padrão”.

16 A fonte destes dados é a página do Haut-Commissariat au Plan, relativos ao Censo de 2014, em <http://rgphentableaux.hcp.ma/>. Estava incluída nos gráficos retirados da página a língua materna hassania, variedade de árabe com influência lexical *amazigh*, de um grupo populacional rural do Sul; não a consideramos aqui, pelo facto de não entrarmos na questão das variedades dialetais do árabe.

17 Criou-se o Instituto Real de Língua e Cultura Amazigh (IRCAM), defendendo uma política virada para uma única versão das três variedades já referidas. Criou-se, também, um canal televisivo *amazigh*, no qual se utiliza esta versão unificada. Na atualidade, nota-se, porém, uma mudança de estratégias: a título de exemplo, na televisão *amazigh*, pode-se escolher uma das variedades para ouvir um programa. Tal opção traz uma certa tranquilidade e satisfação junto da população local.

O panorama sociolinguístico de Marrocos completa-se, em traços largos, com a presença de línguas estrangeiras. Coabitam, com as línguas de matiz nacional, no espaço geográfico marroquino, línguas não originárias do território, mas que podemos considerar veiculares, sobretudo nas zonas urbanas e na comunicação com os estrangeiros, como o francês e, em muito menor grau, o espanhol; são línguas estrangeiras de antiga colonização ou protetorado, cuja implantação se explica, assim, por razões históricas, e se plasma numa repartição geográfica bastante bem definida.¹⁸ Na opinião de Bennis (2011: 2), por exemplo, "as línguas estrangeiras cumprem o papel de línguas de abertura e modernização como no caso do francês e do inglês, enquanto as línguas locais são as línguas de identidade, herança e, em certos casos, da modernidade" (t.n.). Nós poderíamos acrescentar o português.

Marrocos é, assim, um país multilingue, mas é necessário assinalar que esse multilinguismo não é o resultado de uma política linguística estruturada nem de uma opção do governo (Cheddadi, 2011), mas sim resultado das circunstâncias socio-históricas, as quais determinaram a presença de várias línguas no país, verificando-se, para além da diglossia, *lato sensu*, já comentada, a existência de *code-switching*.

O Quadro 3 tenta mostrar as diferentes possibilidades de combinação de línguas, na população marroquina, sendo clara não só a variedade de perfis de conhecimento linguístico encontrados na sociedade, como na nossa Licenciatura:¹⁹

Quadro 3 – Estimativa de combinatórias linguísticas possíveis, num mesmo indivíduo, em Marrocos

Amazigh	Árabe marroquino	Árabe marroquino standard	Árabe clássico	Árabe moderno	Francês	Espanhol
	+		+		(+)	(+)
+			+			
+	+		+		(+)	(+)
	+	+	+	+	(+)	(+)
+	+	+	+	+	(+)	(+)

Assim, resulta difícil identificar, à partida, as línguas utilizadas na sala de aula de PLE, na Licenciatura em Estudos Portugueses, estando os docentes perante turmas que se revelam, geralmente, heterogéneas, o que exige o recurso ao seu próprio conhecimento

18 Marrocos conheceu, desde o fim do séc. XIX, mas sobretudo a partir de 1912, a presença de potências externas: em 1912, foram instituídos o extenso Protetorado francês e o Protetorado espanhol (todo o Norte, excluindo Tânger), à luz de tratados assinados com o sultão Moulay Hafid, os quais duraram até à independência de Marrocos, em 1956. Segundo Dieste e Villanova (2013: nota 4) "O tratado [com Espanha] também estabeleceu outra área de protetorado espanhol no sul do Marrocos, a região de Cap Juby / Tarfaya (...). A ação colonial espanhola foi mínima neste território, mais ligado à colónia do Saara espanhol (Saguia el-Hamra e Río de Oro), que era contígua a ele, do que à zona norte do protetorado". A língua inglesa impõe-se de modo notável no mercado linguístico marroquino, estando incluída no conjunto de línguas estrangeiras do sistema educativo nacional suplantando o francês.

19 Alguns indivíduos falam apenas *amazigh*, nas zonas rurais. Acrescente-se que se parte do princípio de que a maioria da população frequenta a mesquita, onde se usa o árabe clássico, do Corão.

de diferentes línguas faladas em Marrocos, quando, por razões didáticas, desejam fazer comparações linguísticas.

O Projeto MAPEAR²⁰

O Projeto MAPEAR (Construção de materiais de apoio ao ensino / aquisição tardia de Português europeu - língua adicional a / por falantes arabófonos), em curso, visa analisar as interlínguas dos estudantes da Licenciatura, a partir de dados de produção escrita e oral, e chegar a padrões de comportamento linguístico. O seu lançamento visa responder a uma situação enquadrada na história político-social particular de Marrocos e marcada por uma população de estudantes urbana, com um perfil sociolinguístico complexo, como acima tentámos retratar.

Com efeito, e lembrando, os estudantes iniciam o seu contacto com o português e a sua cultura apenas na universidade, tendo 17 anos ou mais (alguns, muito mais), tratando-se, assim, de aprendentes tardios numa língua maioritariamente adicional (L3, L4, Ln); são, maioritariamente, multilingues, sendo que (i) a sua L1 e as restantes variedades nacionais que dominam correspondem a sistema(s) muito afastado(s) do sistema do português e que (ii) muitos deles têm conhecimentos de francês e de espanhol, em grau de proficiência diverso, mas constituindo sempre uma mais-valia para a comparação entre esses sistemas (muito) próximo(s) do português.²¹

Com base nestas premissas, e porque os materiais didáticos disponíveis nem sempre respondem adequadamente às necessidades do público-alvo marroquino, tornou-se clara a necessidade de criar materiais de apoio para docentes e de materiais específicos para os estudantes, a fim de colmatar áreas não adequadamente tratadas nas publicações didáticas disponíveis, face às dificuldades dos estudantes, efetivamente identificadas. Assim, no âmbito do Projeto, prevemos construir conteúdos adaptados à situação específica e pô-los ao serviço dos programas das cadeiras de língua e aos programas a ela associados, incidindo em questões de gramática, tradução, aspetos pragmáticos e cultura portuguesa, *lato sensu*.

Para chegar à delimitação dos domínios mais resistentes à aquisição do PLE, parte-se da análise de interlínguas (considerando como relevante a análise das estratégias usadas pelos estudantes e o papel das línguas que dominam, nacionais e estrangeiras, em particular românicas); como objetivo final, prévio à construção de materiais, interessa-nos, assim, definir padrões de comportamento linguístico. A comparação dos sistemas em presença, a rentabilização de aquisições linguísticas prévias e a pesquisa de informação sobre situações sociolinguísticas análogas fazem igualmente parte da nossa abordagem.

Para trabalharmos com segurança, optámos por constituir um *corpus* de produções

20 O Projeto beneficiou de um subsídio do Camões I.P. para a compra de aparelhos de som e informáticos necessários.

21 Remete-se, para informações complementares, para Pinto (2012).

escritas desenvolvidas em diferentes cadeiras e de produções orais, sob forma de entrevista/conversa livre, acreditando que "grandes *corpora* fornecem muitos factos novos (...) que muitos professores (...) presumiam estarem bastante bem compreendidos" (t.n.).²² A construção do *corpus*, em fase inicial mas em progresso, conta atualmente com 151 produções escritas e 35 minutos de produção oral, de estudantes dos 6 semestres. O formato do *corpus* foi concebido usando os ensinamentos da sociolinguística variacionista, visando ser equilibrado e representativo da população, tendo sido selecionadas diferentes variáveis independentes que nos parecem importantes para o efeito. Dispor de um *corpus* vasto permite evidenciar eventos linguísticos repetidos, recorrentes, chegar a padrões de comportamento e assim orientar as linhas de pesquisa e a construção dos materiais. A recolha é realizada pelos próprios docentes, o que garante a familiaridade com o enquadramento e com o contexto de produção.

O âmbito do trabalho é o dos níveis A1 a B2 do QECR, os quais correspondem aos dois primeiros anos da Licenciatura e ao início do 5º semestre, em termos globais.

Ilustram-se, de seguida, alguns dos domínios que necessitam de particular atenção, no contexto em causa.

Alguns domínios de intervenção e exemplos do *corpus* MAPEAR

A título de exemplo do nosso *corpus* escrito, vejam-se os casos seguintes, com provável interferência de outras línguas românicas ou das L1, que uma análise comparativa alargada, considerando os padrões de comportamento a que se chegar, tentará verificar (sabendo-se, contudo, que existem fases de tentativa ou de testagem de soluções adequadas que podem relevar de estratégias diversificadas). A codificação das produções permite, no momento da análise, acompanhar um mesmo informante do 1º ao 6º semestre, o que é importante para a consideração das fases de aproximação do sistema alvo e dos domínios mais ou menos resistentes à aquisição.

1. Género

O valor de género dos nomes e o valor adquirido pelos adjetivos, em sintaxe, revela-se problemático, como já notado em trabalhos anteriores, como é o caso de Naciri (2012) e Pinto (2017), que apresentam algumas conclusões relevantes e a ter em consideração neste Projeto:

- 1) O professor (...) apresentou um **analise** simple.
- 2) ... e **valores** semanticas que cada palavra tem.

22 Em M. Stubbs, 2004, Language Corpora, cap. 4 de *Handbook of Applied Linguistics*, disponível em https://www.academia.edu/36721987/Handbook_of_Applied_Linguistics

2. Concordância

A concordância no sintagma nominal e no sintagma verbal apresenta-se também problemática, como mostram os exemplos seguintes, sendo que o exemplo 4) revela igualmente a questão da aquisição dos verbos morfologicamente mais complexos:

3) Nós estudantes, vimos **muitas pinturas** francesa e portuguesa.

4) Quando **ela** *estudou* em liceu Descartes tive muitas amigas.

3. Tempos/modos verbais (e *Concordantia/Consecutio Temporum*)

Neste domínio, verifica-se especial dificuldade em distinguir, por um lado, os valores do indicativo perfeito e imperfeito e, por outro lado, em harmonizar indicativo e conjuntivo e formas compostas entre si:

5) No ano passado, **durante as férias**, costumei visitar as bibliotecas.

6) ... embora o termo possa estar errado, **era** o mesmo.

7) Se tiver tempo, **tinha ido** ao cinema contigo.

4. Clíticos

Tratando-se de um domínio particularmente complexo, a aquisição da colocação dos pronomes clíticos permanece uma dificuldade até ao final do percurso:

8) A viagem **se** trata descobrimentos do caminho marítimo de Vasco de Gama.

9) ... tentou analisá-**los** e nos deu uma ideia deles.

5. Fonologia/Ortografia

Verifica-se com frequência a existência de dificuldades na relação fonologia/ortografia, muito provavelmente devido à influência do árabe, afetando léxico com a vogal /e/, em particular:

10) ... uma atitude crítica em relação à igreja.

11) ... para procurar dos produtos e especiarias no outro país.

6. Topónimos e antropónimos/Ortografia

Em árabe, não existem maiúsculas para os chamados nomes próprios, sendo o uso de minúsculas iniciais, em português, um problema que resiste por muito tempo:

12) ... caminho marítimo de vasco de gama.

Estas são apenas algumas das áreas de análise e de intervenção, no âmbito do Projeto, podendo evidenciar-se, através dos poucos exemplos dados, que a questão da interferência das

línguas nacionais e estrangeiras é um aspeto a olhar de perto pela equipa, constituída pelos docentes da Licenciatura de Rabat e por quatro colaboradores portugueses, três dos quais associados à Faculdade de Letras de Lisboa.

Na fase final do Projeto, serão divulgados resultados e o *corpus* será disponibilizado à consulta externa.

Comentários finais e projetos de futuro

A avaliação geral destes doze anos de ativa cooperação institucional é muito positiva, na nossa opinião e como tentámos demonstrar, ao longo desta apresentação. Numa região particularmente importante para a paz e a cooperação na zona mediterrânica ocidental-atlântica, acreditamos que este projeto de formação contribui para o aprofundamento das relações entre países vizinhos, enquanto instrumento «para incentivar as trocas entre os agentes do desenvolvimento: responsáveis da sociedade civil e política, mundo cultural e religioso, universidades, investigadores, meios de comunicação social, associações, sindicatos e empresas públicas e privadas», conforme citação identificada na nota 9 deste texto. O interesse pelo português, em Marrocos, é realmente grande, existindo a vantagem de que Portugal é conhecido pela população em geral como um país amigo.

A coordenação e o corpo docente da Licenciatura, agora autónoma, concordam em continuar a lutar pela obtenção de algumas mais-valias que, desde a fase anterior, se almejava conseguir. Entre elas, destaca-se a inclusão da língua portuguesa como língua estrangeira opcional no ensino secundário marroquino (a nível do *Baccalauréat*), o que permitiria que muitos dos estudantes já dispusessem de conhecimento prévio desta língua, ao entrar no ensino superior.

Seria, igualmente, muito importante obter do Ministério da Educação Marroquino a abertura de cursos de português noutras universidades marroquinas, sendo várias as que nos declararam ter interesse nessa abertura (por exemplo, as Universidades de El-Jadida, de Fez, de Casablanca), em alguns casos para estudantes de cursos técnicos. Essa possibilidade permitiria dar emprego no ensino a pós-graduados da UM5R.²³

Prevista a breve trecho está a inclusão do português no Mestrado em Tradução, recentemente aberto na FLSH de Rabat, o que constitui uma conquista muito relevante, permitindo a posterior formação dos muito necessários tradutores ajuramentados de árabe-português/português-árabe. O intercâmbio académico irá fortalecer-se ainda mais com a nova reforma académica do *Bachelor*, a qual prevê que a Licenciatura em Estudos Portugueses, como as suas congéneres, passe a ser lecionada em 4 anos e com o sistema internacional de créditos, permitindo assim uma maior facilidade na mobilidade internacional.

23 Na sede da Câmara de Comércio Luso-marroquina, em Casablanca, têm vindo a ser dados alguns cursos de português, mas atualmente existem algumas dificuldades logísticas.

Está também prevista a abertura, logo que possível, de um centro de exames do CAPLE²⁴ na Faculdade de Letras de Rabat, o qual seria o primeiro no Mundo Árabe, permitindo aos graduados da UM5R e a estudantes de outras nacionalidades magrebins a obtenção de uma certificação internacionalmente reconhecida em PLE.

Finalmente, e dada a existência dos já referidos protocolos universitários entre Rabat e Lisboa, contamos organizar colóquios sobre temáticas de mútuo interesse e publicar os resultados dos mesmos, logo que a situação sanitária mundial o permita.²⁵

Anexo

Dada a importância do francês e, em parte, do espanhol no território marroquino e da sua integração na mesma família linguística do português, com as consequências já referidas para a aquisição do português, deixamos aqui um complemento de informação.

Como ponto prévio, é de notar que, nos tempos da colonização/do protetorado, Marrocos era dividido em três regiões: região francesa, região espanhola e cidade de Tânger que, pela situação geográfica, era uma região multinacional (Roldán, 2005). A presença da língua francesa nos territórios marroquinos remonta ao início do século XX, mais precisamente ao ano de 1907. A língua francesa era a língua oficial das regiões francesas e usava-se independentemente das línguas autóctones. Nos tempos da independência, esta língua adquiriu o estatuto de primeira língua estrangeira. De facto, a independência trouxe uma nova repartição das línguas no mapa nacional. Krikez (2005, *apud* Roldán, 2005: 38) reparte o modo de considerar a língua francesa nos tempos da independência em três etapas. Na primeira etapa (1956-1970), a da independência, mesmo sendo determinado que a língua oficial do país é o árabe, o francês, pela inexistência de funcionários marroquinos tanto nas administrações como no ensino, continuava a manter o seu estatuto, como nos tempos da colonização. Nos anos oitenta, o francês começou a deixar espaço ao árabe, facto animado pela subida das vozes que reclamavam o perigo daquela língua para a identidade muçulmana do país. Começou então o processo de arabização em todos os domínios, reduzindo-se o francês ao âmbito do ensino científico e técnico. Nessa altura, o conceito fundamental era o de considerar o árabe como língua oficial e o francês como primeira língua estrangeira. A última etapa definida por este investigador vai desde 1980 até à atualidade. Nesta última etapa, consequência das anteriores, a língua francesa é a primeira língua estrangeira do país, língua de abertura ao mundo, que interessa ser ensinada devido às características socioculturais marroquinas²⁶, e que está presente na administração, mesmo no norte marroquino onde se fala mais o espanhol.

24 Centro de Avaliação e de Certificação de Português Língua Estrangeira, Faculdade de Letras de Lisboa.

25 A FLUL tem grande tradição de estudos árabes - língua e cultura, com forte presença no Centro de História da FLUL (arte e património português construído em Marrocos, presença berbere e árabe em Portugal, estudos de diplomacia bilateral, etc.), sem esquecer que o Campo Arqueológico de Mértola é obra de um ex-docente do Departamento de História da FLUL e grande especialista do mundo árabe, Cláudio Torres.

26 De acordo com o jornalista Ali Amar, citado em Albet-mas et al. (1995), "A forte presença da cultura francesa em escolas, colégios e institutos perpetuou (...) [a] reprodução das elites francófilas, que geralmente concluem o seu ensino superior em França" (t.n.).

Devido à proximidade geográfica com Espanha, à movimentação das populações entre os dois países, nomeadamente para trocas comerciais, à presença dos enclaves espanhóis de Ceuta e Mellila e por razões de índole histórica, o espanhol tem uma posição importante na costa norte de Marrocos. Boukous (1995: 22) considera que a presença do espanhol em Marrocos data do século XVI, época em que os mouriscos e os judeus encontraram refúgio nas cidades marroquinas da zona norte. Essa presença reforçou-se no séc. XIX, com focos de administração espanhola imposta pela força e, posteriormente, com a chegada dos colonizadores, no séc. XX (1912, Tratado de Fez), que aí permaneceram por aproximadamente quarenta anos.

Referências

ALBET-MAS, Abel; GARCIA-RAMON, M. Dolors; NOGUE-FONT, Joan; RIUDOR-GORGAS, Lluís. Géographie, aménagement du territoire et colonialisme espagnol au Maroc. *Cahiers de géographie du Québec*, 39(106), p. 43–59, 1995. Disponível em: <https://doi.org/10.7202/022477ar>. Acesso em: novembro de 2021.

BENNIS, Saïd. La situation linguistique au Maroc : Enjeux et état des lieux, 2011. Disponível em: https://www.academia.edu/7281190/Situation_linguistique_au_Maroc. Acesso em: novembro de 2021.

BOUKOUS, Ahmed. *Société, langues et cultures au Maroc : enjeux symboliques*. Rabat: Faculté des Lettres et des Sciences Humaines, 1995.

CHEDDADI, Abdesslam. Al-waḍc al-luġawī al-maġribī bayn 'irṭ al-māḍī wa muqtaḍayāt al-ḥāḍer. *Al-Madrassa Al-Maġhribiya*, n. 3, p. 13-41, 2011.

DIESTE, Josep Lluís Mateo; VILLANOVA, José Luis. Les interventores du protectorat espagnol au Maroc. Contextes de production d'une connaissance politique des cabilas. *Cahiers d'études africaines*, 53 (3), n. 211, p. 595-624, 2013.

ENNAJI, Moha. De la diglossie à la quadriglossie. *Langues et linguistiques, revue internationale de linguistique*, n. 8, 49-64, 2001.

PINTO, Jorge. Transferências lexicais na aquisição de português como língua terceira ou língua adicional. Um estudo com alunos universitários em Marrocos. *Diacrítica*, 26/1, p. 171-187, 2012.

_____. A aquisição do género e da concordância de género em português língua terceira ou língua adicional. *Teorias e Usos Linguísticos*. Lisboa: Lidel, p. 91-110, 2017.

NACIRI, Habiba. O multilinguismo e os processos de aquisição de PLE na universidade marroquina. Tese (Doutoramento), Universidade Mohammed V-Agdal, Rabat, 2012.

ROLDÁN ROMERO, Magdalena. El español en el contexto sociolingüístico marroquí: evolución y perspectivas. *Aljamía*, p. 37-46, 2005.

YOUSSI, Abderrahim. La triglossie dans la typologie linguistique. *La linguistique*, v. 19, p. 71-83, 1983.



VARIAÇÃO LINGUÍSTICA, ENSINO DE LÍNGUA PORTUGUESA E ATITUDES LINGUÍSTICAS LINGUISTIC VARIATION, PORTUGUESE LANGUAGE TEACHING AND LINGUISTIC ATTITUDES

Juliana Bertucci Barbosa¹

Daiana Lombardi de Cuba²

RESUMO

Dentre os grandes desafios que estão presentes no ensino de língua portuguesa nas escolas, um dos mais polêmicos é o trabalho com a variação linguística. Sabemos que a tradição do estudo prescritivo da gramática normativa e a visão de língua homogênea nas escolas ainda é presente, apesar de os documentos oficiais do governo brasileiro orientarem, desde 1997, com os Parâmetros Curriculares Nacionais (PCN), o trabalho escolar sob uma perspectiva de língua dinâmica e variável. Pesando tais considerações, este artigo discorre sobre os novos caminhos da Sociolinguística (Educativa) e apresenta resultados sobre um teste de atitudes linguísticas aplicado a alunos de duas turmas do 7º Ano da Educação Básica de uma escola pública da cidade de Uberaba, revelando o que pensam sobre e como avaliam a língua portuguesa e suas variações.

PALAVRAS-CHAVE: Ensino de língua portuguesa brasileira; atitudes linguísticas; Educação Básica.

ABSTRACT

Among the great challenges that are present in Portuguese language teaching in schools, one of the most controversial is the work with linguistic variations. We know that the tradition of the prescriptive study of normative grammar and the vision of homogeneous language in schools is still present, despite official documents of the Brazilian government guiding, since 1997, with the National Curriculum Parameters (PCN), school work from a perspective of dynamic and variable language. Weighing these considerations, this article discusses new paths in Sociolinguistics (Educational) and presents results on a test of linguistic attitudes designed for students from two classes of the 7th Year of Basic Education of a public school in the city of Uberaba, revealing what they think about and how they evaluate the Portuguese language and its variations.

KEYWORDS: Brazilian Portuguese language teaching; Linguistic attitudes; Basic education.

1 Professora Associada do Departamento de Linguística e Língua Portuguesa e do PROFLETRAS da Universidade Federal do Triângulo Mineiro (UFMT), Sede Uberaba, MG. Professora Colaboradora do Programa de Pós-Graduação em Linguística e Língua Portuguesa UNESP/Araraquara. E-mail: julianabertucci@gmail.com

2 Doutoranda no Programa de Pós-Graduação em Linguística e Língua Portuguesa UNESP/Araraquara, Egressa do PROFLETRAS/UFMT e Professora de língua portuguesa da Educação Básica de Uberaba, MG. E-mail: daiana.lombardi@unesp.br

Introdução³

São muitos os desafios que estão presentes no ensino de língua portuguesa nas escolas. Um deles, e talvez o mais polêmico, é o trabalho com a variação linguística em sala de aula. Sabemos que a tradição do estudo prescritivo da gramática normativa nas escolas ainda é muito forte e que muitos professores passam para os alunos a impressão de que a língua portuguesa é única e que existe apenas uma forma correta de expressar-se. Sabemos também que o trabalho com a variação linguística na escola continua muito tímido e, muitas vezes, deixa de proporcionar aos alunos reflexões sobre o funcionamento da(s) língua(s) e de como o aspecto social interfere sobre este funcionamento.

Mesmo tendo consciência dessa tradição gramatical, o trabalho com a variação linguística em sala de aula não é uma novidade. Documentos oficiais do governo brasileiro que orientam o ensino de língua portuguesa no Brasil, como os Parâmetros Curriculares Nacionais - PCN (BRASIL, 1997), desde o final dos anos 90 apresentam uma perspectiva de língua dinâmica, plural e preveem que o estudo e o (re)conhecimento da variação linguística no/do Português Brasileiro (PB) devem estar presentes na escola. O mais recente documento segue essa mesma linha, a Base Nacional Curricular Comum - BNCC (BRASIL, 2018), e continua também abrindo espaço para o trabalho com a variação e a diversidade linguística brasileira nas aulas de língua portuguesa.

Considerando tais contextualizações, e motivadas por pesquisas realizadas no âmbito do Programa de Pós-Graduação de Mestrado Profissional em Letras (PROFLETRAS) e no Grupo de Pesquisa em Estudos Variacionistas (GEVAR/Diretório CNPq), outras inquietações surgiram, tais como: Os alunos têm conhecimento da existência das diferentes variedades da língua portuguesa? Como eles avaliam essas variedades?

Partindo dessas indagações, a pesquisa apresentada neste artigo visou investigar o que os alunos do 7º Ano de uma escola pública da cidade de Uberaba pensam sobre e como avaliam a língua portuguesa e suas variedades. A partir desta nossa investigação, realizada por meio de testes de atitudes linguísticas, a pesquisa também intenta colaborar com reflexões sobre essas atitudes e, conseqüentemente, sobre o ensino de língua portuguesa na Educação Básica (especificamente, no Ensino Fundamental, ciclo final) levando em conta o tratamento dado à diversidade linguística em sala de aula.

3 Este trabalho foi desenvolvido a partir de investigações realizadas no âmbito de projeto fomento pela CNPq (Processo 424520/2016-8) e pela CAPES.

A Pedagogia da Variação Linguística

A Pedagogia da Variação linguística é uma proposta relativamente nova, que emerge da recente área intitulada Sociolinguística Educacional (BORTONI RICARDO, 2004) e é defendida por autores como Zilles e Faraco (2015), Bagno (2007), Faraco (2008, 2011), Bortoni-Ricardo (2004, 2005), Cyranka (2015), Barbosa e Cuba (20015), Marine e Bertucci Barbosa (2017), entre outros.

Na obra *Pedagogia da variação linguística* (2015), organizada pelos autores Ana Maria Zilles e Carlos Alberto Faraco, encontramos vários capítulos que apresentam pesquisas sobre variação linguística e o ensino, agrupados nos seguintes temas: “Variação e práticas escolares”, “Variação em foco e suas implicações pedagógicas”, “Variação linguística no domínio público” e “Olhares acadêmicos sobre a variação linguística e preconceito”. O livro evidencia o quanto o campo de estudos e as pesquisas nesta área – variação linguística e ensino – é fértil e o quanto há ainda o que ser feito/pesquisado nessa área.

Nas palavras dos organizadores, na introdução do livro:

A língua continua sendo forte elemento de discriminação social, seja no próprio contexto escolar, seja em outros contextos sociais, como no acesso ao emprego e aos serviços públicos em geral (serviços de saúde, por exemplo). Por isso parece ser um grande equívoco a afirmação de que a variação linguística não deve ser matéria de ensino na escola básica. Assim, a questão crucial para nós é saber como tratá-la pedagogicamente, ou seja, como desenvolver uma pedagogia da variação linguística no sistema escolar de uma sociedade que não reconheceu sua complexa cara linguística e, como resultado da profunda divisão socioeconômica que caracterizou historicamente sua formação (...), ainda discrimina fortemente pela língua os grupos socioeconômicos que recebem as menores parcelas da renda nacional. (ZILLES; FARACO, 2015, p.8)

Isso prova o quanto se faz necessário um trabalho efetivo e reflexivo com a variação linguística em sala de aula, com os valores sociais da linguagem, extinguindo-se o preconceito linguístico tão arraigado em nossa sociedade, como enfatizam Zilles e Faraco (2015).

Cyranka (2015), em um dos capítulos do referido livro, expõe o complexo contexto escolar em que vivemos, ponderando que precisamos

investir no grande desafio posto no contexto escolar brasileiro, principalmente nas últimas décadas do século XX e que vem se agravando neste começo do século XXI. O fracasso escolar se agiganta entre nós. Misturam-se propostas teóricas avançadas, investimentos dos governos federal, estadual e municipal na construção de um discurso de combate ao analfabetismo funcional; sempre novos projetos, novas propostas. Continuamos, no entanto, preocupados com a defasagem entre o que é necessário saber na sociedade contemporânea, dominada pela escrita, e o que nossas crianças e nossos jovens têm demonstrado alcançar em competência de leitura e de escrita. Sem dúvida, a sociolinguística educacional, a pedagogia da variação linguística têm um importante papel a desempenhar nessa área do ensino. (CYRANKA, 2015, p. 33)

A autora também esclarece, em outro capítulo, do livro *Sociolinguística, Sociolinguísticas* (2016), que:

A Sociolinguística Educacional propõe que se leve para as salas de aula a discussão sobre variação linguística, orientando os alunos a reconhecerem as diferenças dialetais e, mais importante, a compreenderem que essas diferenças são normais, legítimas e que devem ser consideradas na seleção das estruturas a serem utilizadas, a depender das condições de produção, isto é, das necessidades do leitor/escritor, falante/interlocutor, a partir do contexto em que se encontra. Desse modo, o conceito de “certo/errado” em linguagem é substituído pelo de “adequado/inadequado”, o que predispõe os alunos ao desejo de ampliarem a competência comunicativa que já possuem, construindo crenças positivas sobre o conhecimento que têm de sua língua, no caso, a língua portuguesa. Crenças positivas levam a atitudes linguísticas positivas, o que garante boa autoestima e entusiasmo do aluno. (CYRANKA, 2016, p. 169-170)

Por meio dessas palavras de Cyranka (2015, 2016), percebemos o quanto é importante discutir em sala de aula questões relacionadas à variação linguística para que eles construam crenças e atitudes linguísticas positivas sobre as variedades da língua, elevando assim a autoestima, entusiasmo e, conseqüentemente, o desempenho linguístico dos alunos. A autora, ainda nesse capítulo, argumenta que é preciso levar para as salas de aula, sem receios, a reflexão sociolinguística para que os alunos reconheçam a legitimidade de todas as variantes utilizadas entre os falantes juntamente com a necessidade de se adequá-las às condições de produção. A autora ainda deixa claro que é papel da escola também implementar práticas de leitura e escrita que favoreçam o contato dos alunos com diferentes gêneros e com as variedades cultas da língua portuguesa, para, dessa forma, levar os alunos a se tornarem competentes também nos usos dessas variedades.

É importante ressaltar que, muitas vezes, o professor de língua portuguesa sente-se despreparado para o trabalho com a variação linguística em sala de aula, seja por lacunas em sua formação acadêmico-profissional; seja por insegurança em abordar o tema, preferindo trabalhar com atividades de leitura e escrita (que também são importantes) mais “tradicionais” que não levam os alunos a pensarem/refletirem/pesquisarem sobre a língua portuguesa. Considerando isso, Cuba (2019) elaborou e organizou além de um Caderno de Atividades voltado para anos do Ensino Fundamental (ciclos finais) sobre os diferentes tipos de variações linguísticas a partir de diversos gêneros textuais, um caderno com orientações aos professores para que eles se sentissem mais seguros para trabalhar com a variação linguística em sala de aula.

Cyranka (2016, p. 154), em relação à formação docente, ainda afirma:

(...) os professores precisam estar sustentados por uma reflexão teórica suficiente. Precisam se tornar também pesquisadores e aprender a encontrar soluções novas a partir do que propõem as teorias linguísticas. Devem ser sensíveis à questão do perigoso distanciamento entre o padrão escolar e a realidade sociocultural de seus alunos, estando aí incluídos seu dialeto e o de sua comunidade linguística. Em vez de investir na valorização pura da norma-

padrão, reconhecer nela a ideologia pela qual, certamente, estarão também afetados. Devem ser capazes de responder à tradição e trilhar pelos caminhos novos que lhes são oferecidos pelas perspectivas abertas pela Sociolinguística Educacional. Precisam reconhecer que seus alunos são usuários competentes de um sistema complexo que eles próprios adquiriram naturalmente e com o qual são capazes de atuar em seu meio social. À escola cabe levá-los a ampliar esses recursos.

A pesquisadora deixa claro também, em suas considerações finais do capítulo, que o fato de a escola reconhecer a legitimidade do vernáculo utilizado por quem quer que seja não a deve impedir de permitir ao aluno o acesso ao domínio das variedades linguísticas que ele desconhece, como, por exemplo, as variedades cultas da língua.

Outra grande sociolinguista, que propôs a denominação do termo Sociolinguística Educacional, é Bortoni-Ricardo (2004, 2005) que, em suas obras, destaca a importância da escola como “uma das principais instituições que influenciam o curso das línguas na sociedade” (BORTONI-RICARDO, 2005, p. 153). Para a autora, a escola é a principal instituição de perpetuação do poder, por isso, urge que utilizemos esse poder para romper com o preconceito linguístico e promover uma pedagogia da variação linguística na qual nossos alunos (e professores) se conscientizem de que:

(...) todas as variedades que compõem a ecologia linguística de uma comunidade, sejam elas línguas distintas ou dialetos de uma ou de mais de uma língua, são funcionalmente comparáveis e essencialmente equivalentes. Nenhum deles é inerentemente inferior, e, portanto, seus falantes não podem ser considerados linguística ou culturalmente deficientes. (BORTONI-RICARDO, 2005, p. 151)

Faraco (2007, 2011), grande defensor da pedagogia da variação linguística, propõe que os professores expliquem aos alunos que não existe erro nos usos que os falantes fazem de sua língua e que não consideremos como “erro” as variedades linguísticas que os alunos trazem para a sala de aula, cabendo ao professor lhes ensinar as variedades cultas para que os alunos possam ampliar seu repertório linguístico.

Bagno (2007, p. 82) também propõe que haja uma reeducação sociolinguística, na qual o professor use o espaço e o tempo escolares para formar cidadãos conscientes da complexidade da dinâmica social e das múltiplas escalas de valores que empregamos a todo momento em nossas relações com os outros por meio da linguagem. O autor também propõe, que:

(...) ao lado do desenvolvimento da leitura e da escrita - prioritário -, também é importante levar o aluno a refletir sobre a língua, a se deter no exame das regras que fazem a língua funcionar como funciona (...). Afinal, saber coisas sobre a língua faz parte da nossa tradição cultural. Muito do que aprendemos na escola não tem uma utilidade prática na vida diária, mas contribui para a nossa formação intelectual, para nossa inserção na cultura mais ampla da qual fazemos parte. (BAGNO, 2007, p. 194-195)

Sendo assim, compartilhando das concepções dos autores aqui citados, também defendemos que o trabalho com a língua em sala de aula deve proporcionar aos alunos uma reflexão sobre o funcionamento das línguas, sobre o porquê das variações, como as mudanças ocorrem, quais fatores interferem nessas mudanças e, ainda, quais as consequências dessas mudanças na sociedade, realizando um trabalho não apenas receptivo de normas e regras, mas, como propõe Bagno (2007), que seja feito um ensino reflexivo-crítico-investigativo em sala de aula.

Marine e Bertucci Barbosa (2017) também defendem, em seu artigo “Em busca de um Ensino Sociolinguístico de Língua Portuguesa no Brasil”, uma pedagogia mais sensível à variação linguística e destacam as contribuições da Sociolinguística ao ensino de língua portuguesa citadas a seguir:

- Respaldo teórico-metodológico para compreender a variação linguística como um fato da língua e não como um problema, uma deteriorização ou “perigo”;
- Definição de conceitos básicos para o tratamento adequado dos fenômenos variáveis, pondo um ponto final à nociva e equivocada “doutrina do erro”;
- Reconhecimento da pluralidade linguística brasileira formada por diferentes línguas (autóctones, alóctones, LIBRAS) e variedades de língua em convivência no seio social;
- Identificação de crenças e atitudes linguísticas diretamente ligadas ao ensino-aprendizado da língua portuguesa;
- Entendimento do “prestígio encoberto” (LABOV, 1972) e suas implicações ao ensino da norma culta;
- Compreensão dos conceitos de norma, norma culta (concreta e plural), norma padrão (idealizada e singular)¹⁷;
- Compreensão dos conceitos de variedades cultas e variedades populares;
- Conceituação de registro e modalidade (instâncias distintas);
- Reconhecimento de que a fala e a escrita apresentam normas diferentes e se realizam num *continuum* de formalidade (MARCUSCHI, 2007) ou dialetal (BORTONI-RICARDO, 2005);
- Reconhecimento das semelhanças entre as variedades cultas e populares;
- Pesquisas sociolinguísticas de caráter descritivo que nos possibilitam (re) conhecer a língua em uso, apontando para uma configuração – considerando-se diferentes fenômenos em diversos níveis linguísticos – bastante distinta daquela prescrita pelas gramáticas normativas escolares e pelos livros didáticos. (MARINE, BERTUCCI BARBOSA, 2017, p. 209-210)

Como vimos, partindo dos apontamentos dos autores citados, são muitas e grandes as contribuições da Sociolinguística para o ensino de língua portuguesa nas escolas. A grande questão é como e quando essas contribuições chegarão a nossas escolas e serão utilizadas efetivamente em nossas salas de aula através de atividades que proporcionem uma maior reflexão da língua e de seus usos nas mais variadas situações comunicativas (orais e escritas). Faz-se necessário, então, levar aos nossos alunos uma questão fundamental, pregada pela pedagogia da variação linguística, de que não há variedades linguísticas inferiores a outras e que não existe maneira “certa” ou “errada” de falar, mas sim, há diversidades sociais e culturais e situações de maior ou menor grau de monitoramento estilístico em que se torna adequado ou não o uso de determinadas variedades em detrimento de outras.

Como vimos, por tratar de uma área com pesquisas ainda promissoras, há muito o que ser feito para que possamos ter nas salas de aula uma pedagogia da variação linguística. Sobre isso, retomando o que nos aponta Faraco (2011, p.8), se “como resultado da intervenção dos linguistas, o tema da variação acabou incorporado pelo discurso pedagógico, podemos dizer que não conseguimos ainda construir uma pedagogia adequada a essa área.”. Sendo assim, esta investigação procura dar uma pequena parcela de contribuição para que a variação linguística possa ser tema de estudo/reflexão em sala de aula.

Variação linguística, documentos oficiais e materiais didáticos

Sabemos que o trabalho com a variação linguística nas escolas ainda é pouco presente. Muitos professores sentem-se inseguros e despreparados para abordar esse tema nas aulas de língua portuguesa e, ainda, os próprios livros didáticos, material de uso de professores e alunos das redes públicas, abordam (quando abordam) muito pouco (e às vezes equivocadamente) sobre as variedades do português. As provas elaboradas pelo governo também apresentam dificuldades terminológicas ao tratar do tema. Pesquisas como as empreendidas por Bagno (2007), Bortoni-Ricardo (2005), Lima (2016), González (2015), entre outros, exemplificam isso.

Apesar de a variação linguística ser pouco trabalhada na escola, em um documento oficial que norteou por algumas décadas o ensino de língua portuguesa no país – os Parâmetros Curriculares (PCN) de Língua Portuguesa –, desde o ano de 1997 (há mais de 20 anos), já nos orientava que:

A Língua Portuguesa, no Brasil, possui muitas variedades dialetais. Identificam-se geográfica e socialmente as pessoas pela forma como falam. Mas há muitos preconceitos decorrentes do **valor social relativo que é atribuído aos diferentes modos de falar:** é muito comum se considerarem as variedades lingüísticas de menor prestígio como inferiores ou erradas. **O problema do preconceito disseminado na sociedade em relação às falas dialetais deve ser enfrentado, na escola, como parte do objetivo educacional mais amplo de educação para o respeito à diferença.** Para

isso, e também para poder ensinar Língua Portuguesa, a escola precisa livrar-se de alguns mitos: o de que existe uma única forma “certa” de falar — a que se parece com a escrita — e o de que a escrita é o espelho da fala — e, sendo assim, seria preciso “consertar” a fala do aluno para evitar que ele escreva errado. Essas duas crenças produziram uma prática de mutilação cultural que, além de desvalorizar a forma de falar do aluno, tratando sua comunidade como se fosse formada por incapazes, denota desconhecimento de que a escrita de uma língua não corresponde inteiramente a nenhum de seus dialetos, por mais prestígio que um deles tenha em um dado momento histórico. **A questão não é falar certo ou errado, mas saber qual forma de fala utilizar, considerando as características do contexto de comunicação**, ou seja, saber adequar o registro às diferentes situações comunicativas. É saber coordenar satisfatoriamente o que falar e como fazê-lo, considerando a quem e por que se diz determinada coisa. É saber, portanto, quais variedades e registros da língua oral são pertinentes em função da intenção comunicativa, do contexto e dos interlocutores a quem o texto se dirige. **A questão não é de correção da forma, mas de sua adequação às circunstâncias de uso, ou seja, de utilização eficaz da linguagem: falar bem é falar adequadamente, é produzir o efeito pretendido.** (BRASIL, 1997, p. 21, grifo nosso)

Nesse fragmento dos PCN de Língua Portuguesa, podemos perceber que o documento do governo já adotava uma noção de língua variável, reconhecia as variedades da língua portuguesa e, ainda, assumia ser papel da escola enfrentar o preconceito linguístico, partindo da concepção de que não existe maneira certa ou errada de expressar-se, mas sim maneiras adequadas a diversas situações comunicativas orais ou escritas em que o falante se encontra. Sendo assim, não faz sentido que a escola, que as aulas de língua portuguesa, continuem trabalhando com uma noção de língua homogênea, em que há apenas uma única maneira “correta” de falar.

Também faz parte dos PCN a orientação de que o ensino de língua portuguesa deve tornar os alunos capazes de “conhecer e respeitar as diferentes variedades linguísticas do português falado;” e “conhecer e analisar criticamente os usos da língua como veículo de valores e preconceitos de classe, credo, gênero ou etnia.” (BRASIL, 1997, p. 33). Nesse sentido nos perguntamos: O ensino de língua portuguesa nas escolas segue essa orientação? É trabalhado na escola o respeito às diferentes variedades da língua? Os alunos (re)conhecem e respeitam as diferentes variedades da língua portuguesa? Para responder a tais perguntas esta pesquisa visou identificar, por meio das atitudes linguísticas de alunos de uma escola pública, se as aulas de língua portuguesa colaboram para um ensino de língua mais crítico-reflexivo sobre a variação linguística ou se ainda precisamos enfrentar e acabar com as falsas crenças disseminadas, como a “de que existe uma única forma “certa” de falar”, como mostra o trecho acima retirado dos PCN.

Mais recentemente, temos a Base Nacional Comum Curricular (BNCC) (BRASIL, 2018), que não representa o currículo em si, pois sua finalidade é “orientar os sistemas na elaboração de suas propostas curriculares” (BRASIL, 2016, p.24), respeitando as diversidades (sejam elas de origem social, histórica, cultural) dos alunos. É por considerar essa pluralidade que a BNCC abre o espaço para discussões no campo da diversidade linguística, uma vez

que o documento é contra quaisquer tipos de discriminação (social, regional, gênero, etc.), dentre os quais observamos aquela ocasionada em razão dos usos linguísticos, ou seja, o preconceito linguístico.

Outro documento do governo federal diz respeito aos editais do Programa Nacional do Livro e do Material Didático (PNLD⁴), que apresentam parâmetros a serem seguidos na escolha de materiais didáticos, para que estes possam ser usado nas escolas públicas do país. Sobre os livros didáticos escolhidos ou não por esse programa, Lima (2016) faz uma interessante pesquisa ao analisar o tema da variação linguística nesses materiais recomendados pelo Ministério da Educação e Cultura (MEC) por meio do PNLD.

No capítulo “Variação linguística e os livros didáticos de português”, presente no livro *Ensino de português e sociolinguística*, Lima (2016, p. 117) considera que:

Trazer à tona, em pleno século XXI, o tema da variação linguística para explicá-lo e justificá-lo deveria causar profundas lamentações em todos os que estudam a Sociolinguística e conhecem sua história. Afinal de contas, já são mais de 50 anos de história desse tema e reconhecer que ele é desconhecido de alunos, talvez por negligência de alguns professores de português, nos faz admitir que de algum modo fracassamos em comunicá-lo à comunidade escolar. Nos livros/materiais/manuais didáticos, não há necessidade de se justificar temas como sintaxe, texto, leitura etc., no entanto, parece que precisamos fazê-lo quanto a variação linguística. E isso faz com que esse tema seja visto como mais um conteúdo que deva ter seu próprio capítulo, quando já sabemos que a variação linguística deve estar inserida no ensino dos temas referidos anteriormente.

Em sua pesquisa, Lima (2016) analisa 24 coleções de livros didáticos indicados no “Guia de livros didáticos” (BRASIL, 2009), considerando como são abordados os temas da variação e do preconceito linguísticos. O autor considera o fato de os temas da Sociolinguística estarem incorporados ao livro didático um avanço, mas critica o fato de ser uma incorporação muito superficial, com a finalidade de apenas atender a exigências dos documentos oficiais. O autor mostra, ainda, que em relação ao preconceito linguístico, das 24 coleções analisadas, quatro

4 Programa Nacional do Livro e do Material Didático (PNLD), conforme descrito no site do MEC (<http://portal.mec.gov.br/component/content/article?id=12391:pnld>), é destinado a avaliar e a disponibilizar obras didáticas, pedagógicas e literárias, entre outros materiais de apoio à prática educativa, de forma sistemática, regular e gratuita, às escolas públicas de educação básica das redes federal, estaduais, municipais e distrital e também às instituições de educação infantil comunitárias, confessionais ou filantrópicas sem fins lucrativos e conveniadas com o Poder Público. O Decreto nº 9.099, de 18 de julho de 2017, unificou as ações de aquisição e distribuição de livros didáticos e literários, anteriormente contempladas pelo Programa Nacional do Livro Didático (PNLD) e pelo Programa Nacional Biblioteca da Escola (PNBE). Com nova nomenclatura, o Programa Nacional do Livro e do Material Didático – PNLD também teve seu escopo ampliado com a possibilidade de inclusão de outros materiais de apoio à prática educativa para além das obras didáticas e literárias: obras pedagógicas, softwares e jogos educacionais, materiais de reforço e correção de fluxo, materiais de formação e materiais destinados à gestão escolar, entre outros.

abordam o tema. E dessas, somente duas são analisadas de forma favorável e as outras duas apontam que o trabalho com o tema deve ser feito pelo professor.

Outro estudo que merece nossa consideração é o empreendido por González (2015) que se propôs a analisar a variação linguística em livros de português para o ensino médio. O pesquisador conclui que:

(...) ao contrário do que se viu com relação aos livros didáticos mais adotados, os menos adotados fazem referência à variação linguística em todos os níveis linguísticos: lexical, morfossintático e fonético-fonológico, com variados exemplos. Essa é uma questão muito significativa. A escolha dos livros mais adotados por não apontar a variação em todos os níveis linguísticos pode refletir o entendimento de que o alcance da variação linguística é limitado, acontecendo apenas no nível lexical e, preferencialmente, nas dimensões diatópica e diacrônica. Dessa maneira, se interpretariam os fenômenos variáveis do nível morfossintático como erros e os do nível fonético-fonológico como “sotaque”. No mesmo sentido, ao fazerem a opção por não dar relevo à variação linguística na dimensão diastrática, os livros didáticos mais adotados acabam por não colocar em questão o preconceito linguístico e a violência simbólica que se produz por meio da língua. (GONZÁLEZ, 2015, p. 240)

González (2015, p. 243) também sinaliza, em suas pesquisas, que os livros didáticos mais adotados ignoram a contribuição da variação linguística para a construção de sentidos e expressão de valores sociais, desconsideram o conceito de adequação e quando o consideram é apenas a respeito da “situação de enunciação” e acreditam que a norma culta equivale às prescrições tradicionais, justificando, assim, o uso das regras da gramática normativa como objeto de ensino. O autor evidencia, por meio de seu estudo, que os livros didáticos mais adotados pelas escolas são os mais conservadores e apenas incorporaram (superficialmente) um novo conteúdo, a variação linguística. González chega ao final de sua pesquisa com uma consideração muito válida para a nossa pesquisa, no sentido de mostrar a importância da variação linguística e contribuições que o estudo dela na escola podem oferecer. Nas palavras do autor:

Livros didáticos que se quiserem informativos e produtivos não devem deixar de pensar a variação linguística positivamente. É necessário refletir sistematicamente sobre a variação, discutir suas contribuições para a construção de sentidos, percebê-la atuando em todos os níveis linguísticos e em todas as interações, entendê-la como sinal de riqueza da língua. Muito disso ainda está por fazer e, certamente, não há caminhos dados a percorrer. Os caminhos ainda estão por se construir, e o livro didático pode ajudar nisso. (GONZÁLEZ, 2015, p. 245)

Como podemos observar a partir da pesquisa empreendida por Lima (2016) e González (2015), há ainda muito que fazer para que a variação linguística e tudo que a engloba (tipos de variedades, preconceito linguístico, por exemplo) estejam de fato presentes na escola e, conseqüentemente, no material didático dos alunos. Ficamos com a conclusão feita por Lima

(2016, p. 130) de que a variação linguística “trata-se de um tema que tem seu lugar na formação cultural e social do indivíduo e, portanto, merece ser abordado com profundidade”.

Atitudes linguísticas e ensino de língua: breve conceitualização

Antes de falarmos a respeito das atitudes linguísticas e a importância de levá-las em consideração no ensino de língua portuguesa, cabe mencionar alguns caminhos trilhados pela Sociolinguística, para justificarmos, de certa forma, o interesse pelo trabalho com essa área. As pesquisas sociolinguísticas, nas últimas décadas, têm considerado, cada vez mais, o significado social da variação. Trabalhos da área a respeito de identidades, crenças e atitudes linguísticas dos falantes estão muito frequentes.

Mendes (2017), no capítulo “A terceira onda da Sociolinguística” do livro *Novos caminhos da Linguística*, apresenta várias pesquisas que justificam esse novo caminho que a Sociolinguística vem tomando. O autor explica que:

A terceira onda da Sociolinguística (...) representa isto: uma volta ao significado social da variação. Uma volta, pois, embora os interesses por fatos dessa natureza não seja algo novo (...), pode-se dizer que, rapidamente, na história da disciplina, o significado social da variação cedeu centralidade para o interesse pela mudança linguística. (MENDES, 2017, p.103)

O autor relembra a origem do termo “terceira onda”, com base em Eckert (2012), e também explica que um dos conceitos fundamentais da terceira onda é o da indicialidade dos elementos linguísticos para apontar algum significado social em uma determinada interação entre falantes. Aponta ainda que os estudos dessa onda têm caminhado para experimentos que estimulam reações inconscientes de ouvintes perante diferentes variantes linguísticas. Mendes (2017, p. 116) também apresenta diferentes trabalhos de pesquisadores brasileiros que fazem parte dessa terceira onda da Sociolinguística por considerar as atitudes linguísticas de falantes a respeito de determinadas variedades linguísticas. O autor conclui dizendo que os avanços nos estudos da terceira onda podem auxiliar no entendimento de processos de mudança linguística.

Esta pesquisa, por escolher trabalhar com atitudes/avaliações linguísticas de alunos de uma escola pública da cidade de Uberaba, insere-se nesse contexto atual de estudos/pesquisas da terceira onda da Sociolinguística. Portanto, faz-se necessário entendermos do que se tratam as atitudes linguísticas.

As atitudes, como já dissemos, estão relacionadas à avaliação, à reação de um determinado indivíduo perante algo ou alguma situação que presencia e não apenas com a sua forma de pensar, de ver e de perceber o mundo, como nos explicam Barcelos e Abrahão (2006). Ao fazer a distinção entre crenças e atitudes, os autores demonstram que, embora diretamente relacionadas entre si, cada um desses conceitos apresenta aspectos peculiares. Em relação à

definição de atitudes linguísticas, compartilhamos da definição proposta por Aguilera (2008), em sua pesquisa sobre crenças e atitudes linguísticas dos falantes das capitais brasileiras. Ao fazer referência aos trabalhos de Lambert (1960) e de Moreno Fernández (1998), a autora argumenta que:

(...) a atitude se constitui de três componentes colocados no mesmo nível: o saber ou crença (componente cognoscitivo); a valoração (componente afetivo); e a conduta (componente conativo), 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)

Sem conceituar de forma absoluta o que seria atitude e atitude linguística, podemos dizer que as atitudes podem ser entendidas como a expressão das crenças e dos conhecimentos que um indivíduo tem sobre determinado assunto e que o possibilita avaliar, comportar-se e manifestar sua opinião e sentimentos perante uma situação. As atitudes linguísticas revelam as avaliações dos indivíduos perante uma língua, uma variedade linguística ou uma situação sociolinguística.

Assim, consideraremos, neste trabalho, atitude linguística como uma avaliação (negativa/positiva) do falante em relação a algum dado linguístico, no caso da pesquisa, em relação a diferentes variedades linguísticas, incluindo a de falantes da região (Uberaba-MG).

A respeito disso, Cyranka (2016, p. 143), importante pesquisadora da área de crenças e atitudes linguísticas, esclarece que o “estudo da avaliação implica o das atitudes, isto é, os julgamentos subjetivos dos falantes sobre o uso que eles próprios e seus interlocutores fazem de sua língua, ou mais especificamente, da variedade linguística que utilizam.”. A autora também indica uma grave carência da escola em possibilitar que os alunos adquiram atitudes positivas em relação a sua própria competência de aprender variedades prestigiadas. Cyranka nos mostra que:

Falta, na escola, uma pedagogia adequada ao desenvolvimento de práticas de letramento que, realmente, possibilitem aos alunos desenvolverem competência de leitura e escrita nas variedades cultas da língua. Essa carência resulta nas atitudes negativas em relação a sua própria competência de aprender essas variedades prestigiadas, problema ligado a avaliação, uma das tarefas que a Sociolinguística se propõe (...). (CYRANKA, 2016, p. 143)

A autora, no mesmo artigo em que foi extraída a última citação, ressalta que o trabalho com as atitudes linguísticas relacionado ao ensino de língua é um assunto atual e pertinente. Sobre isso a autora afirma que:

A avaliação linguística é um dos itens que precisa passar a ter destaque nas discussões contemporâneas sobre o trabalho escolar com a língua materna, já que ela desencadeia o processo de construção de julgamentos subjetivos do falante em relação a sua língua, a seu dialeto e ao de seu interlocutor na construção das chamadas atitudes linguísticas. (CYRANKA, 2016, p. 133)

Botassini (2015) enfatiza o quanto pesquisas envolvendo crenças e atitudes linguísticas são importantes e apresenta várias contribuições que esses estudos podem proporcionar.

Como vimos, o trabalho com atitudes linguísticas pode ser muito útil no entendimento de como os falantes consideram/avaliam sua variedade e a dos outros. Corbari (2013, p. 61) em sua tese de doutorado, mostra que o conceito de atitude linguística pode englobar diversas dimensões “desde as atitudes com relação a variedades linguísticas/dialetais e estilos de fala, passando pelas atitudes com relação ao aprendizado de uma língua, até as atitudes com relação a grupos, comunidades, minorias”.

Sendo assim, ao levar pesquisas a respeito das atitudes linguísticas para a escola, para relacionarmos as contribuições dessas pesquisas ao ensino, poderemos contribuir com reflexões acerca de como os alunos lidam com (e se reconhecem) as variações/variedades linguísticas. Como já dissemos na seção deste artigo sobre Pedagogia da Variação linguística, considerar/saber o que os alunos das escolas públicas pensam sobre a língua que aprendem, sobre a língua que falam é muito importante, haja vista que suas avaliações sobre a língua portuguesa e suas variedades influenciam fortemente (acreditamos) no processo de ensino-aprendizagem. Além disso, conhecer suas atitudes possibilita-nos compreender o porquê da dificuldade em usarem a própria língua para se expressarem bem, tanto na oralidade quanto na escrita, e também para nortear o trabalho do professor no sentido de incentivar um ensino de língua portuguesa mais reflexivo-crítico e de erradicar o preconceito linguístico tão presente na escola e fora dela.

Marques e Baronas (2015) apresentam uma pesquisa semelhante à nossa, ao considerarem crenças e atitudes linguísticas de alunos do ensino fundamental de duas escolas públicas de Londrina. As pesquisadoras nos mostram que pesquisar sobre as atitudes linguísticas no âmbito escolar é importante porque estas:

(...) como já sugeriu Morales (1993), afetam não somente fenômenos particulares e específicos, mas principalmente o ensino de línguas maternas e estrangeiras. Com relação à variação linguística, portanto, destacamos que a construção e a abordagem de uma pedagogia diferenciada na escola dependem de crenças e atitudes que se constroem e fomentam a respeito das variedades da língua em uso. (MARQUES E BARONAS, 2015, p. 1)

Sobre essa questão do ensino, mais especificamente sobre o papel do professor, as pesquisadoras Marine e Bertucci Barbosa (2017, p. 195) defendem que:

(...) se os professores passarem a ter uma abordagem mais reflexiva e embasada pelos preceitos da Sociolinguística Educacional, poderemos modificar as crenças dos alunos em relação à língua e, assim, as atitudes em relação às suas próprias variedades linguísticas poderão ser alteradas também. Por isso, o professor de língua portuguesa deve levar em consideração as crenças linguísticas de seus alunos em suas aulas e, a partir desse (re)conhecimento, propor atividades que reflitam sobre o uso da língua nos mais diversos contextos de uso (...).

Os professores, como propagadores de crenças e construtores de atitudes, têm grande influência na construção de crenças e de atitudes dos alunos. Botassini (2015) explica que:

Muitas vezes, a incapacidade de o indivíduo compreender por qual motivo ele sente e reage de uma maneira específica diante de um objeto social o leva a observar os pensamentos e as crenças alheias, podendo adotá-los para justificar seus próprios sentimentos e tendências reativas. Observa-se, então, nesse caso, o princípio de transferência. É por meio da observação, da identificação e da imitação de “professores” sociais (pais, familiares, amigos, professores, figuras públicas...) que se aprendem e se formam as atitudes. (BOTASSINI, 2015, p. 118)

Sendo assim, trabalhar com atitudes linguísticas no ensino de língua materna se faz necessário para que possam ser construídas atitudes positivas em relação à língua e para que os alunos possam: ter uma autoestima elevada sobre a língua (variedade) que falam; ter interesse pelas aulas de língua portuguesa; se desempenharem melhor nas diversas situações em que precisem se expressar linguisticamente e não julgar negativamente a maneira de falar de ninguém. Bagno (2007, p. 207) também deixa claro que as “atitudes da professora em sala de aula, no tratamento dado aos fenômenos de variação linguística, podem exercer uma grande influência no comportamento de seus alunos”. O sociolinguista, ainda sobre as atitudes linguísticas negativas a respeito de determinadas variedades da língua, também considera “inadmissível, nos dias de hoje, que o modo de falar de uma pessoa continue sendo usado como justificativa para atitudes preconceituosas e humilhantes.” (BAGNO, 2007, p. 207). Tal a importância, como vimos anteriormente, possibilita a implementação de uma pedagogia da variação linguística nas escolas para diminuir atitudes negativas para determinadas variedades linguísticas.

O estudo das atitudes linguísticas, apesar de ser um campo relativamente pouco explorado no Brasil, apresenta já relevantes pesquisas a respeito das atitudes linguísticas de falantes no/do território brasileiro, No entanto, ainda há poucos estudos desse campo teórico voltados para o contexto escolar/educacional, visando, principalmente, contribuições para o ensino de língua e formação de professores da Educação Básica. Nesse sentido, nossa pesquisa sobre as atitudes linguísticas de alunos do 7º Ano de uma escola pública de Uberaba pretende contribuir para esse campo de estudos, procurando evidenciar como esses alunos avaliam variedades linguísticas diferentes e refletir sobre a importância de um ensino de língua portuguesa mais crítico-reflexivo-investigativo (BAGNO, 2007) e mais sensível à variação linguística.

Local e informantes da pesquisa

Os informantes da pesquisa são os alunos de duas turmas do 7º Ano de uma escola pública e periférica da cidade de Uberaba. A localização periférica da escola pode ser melhor observada na imagem abaixo:

Figura 1: Distância entre a escola e o centro da cidade



Fonte: Google Maps – adaptado pelas autoras

A escola conta com turmas de Ensino Fundamental (matutino e vespertino) e Médio (matutino) e também oferece Educação em Tempo Integral (ETI) para os alunos do Ensino fundamental. Tem uma infraestrutura adequada ao público escolar e atende alunos da região (periférica) em que se localiza e de regiões próximas. Além disso, tem um Índice de Desenvolvimento da Educação Básica (IDEB) relativamente baixo para o 9º Ano do ensino fundamental: 3.3 (dados do IDEB de 2017).

De maneira geral, os alunos que participaram da pesquisa são da faixa etária entre 12 e 14 anos, de baixa renda, na maioria, uberabenses e moradores do bairro onde se localiza a escola (região periférica da cidade). Nem todos os alunos das turmas participaram da pesquisa, seja por faltas nas aulas ou por sentirem-se inseguros em participar. O número total de alunos dessas duas turmas que participou das três partes do teste e que trouxe a autorização assinada pelo responsável foi de 38 alunos.

Para melhor conhecermos o perfil social e econômico dos alunos participantes da pesquisa, na primeira parte do teste de atitudes os alunos responderam questões sobre isso.

O teste de atitudes e aplicação

Para analisar as atitudes linguísticas desses informantes, elaboramos um teste de atitudes dividido em três partes. A primeira parte é constituída de um questionário socioeconômico e de costumes para conhecermos a realidade do aluno participante da pesquisa, afinal partimos da ideia de que língua e sociedade se relacionam intimamente. Os alunos responderam tal questionário com perguntas de múltipla escolha. A segunda parte do teste foi composta por questões em que

apresentam diferentes variedades linguísticas para os alunos avaliarem. As perguntas dessa parte do questionário eram abertas e os alunos ficaram livres para exporem suas opiniões.

A última parte do teste foi composta por quatro pequenas gravações, cada gravação com um falante diferente (idade, sexo, região de origem, escolaridade), as quais foram avaliadas pelos alunos, seguindo o modelo de teste de atitudes utilizado por Cyranka (2007). Os alunos avaliaram o falante, a partir da escuta do áudios, quanto aos graus de: inteligência, poder aquisitivo, honestidade, simpatia, competência no trabalho/estudo e virtude da pessoa (boa, má, etc), seguindo uma escala de notas de 1 a 7. É importante ressaltar que, nas gravações, os informantes relataram fatos da própria infância.

Para a aplicação do teste, após o consentimento da direção da escola, foi explicado aos alunos de cada turma do 7º Ano da escola do que se tratava a pesquisa e quais eram os objetivos. Foi-lhes esclarecido também o quanto sua participação era importante e que em momento algum teriam suas identidades expostas na pesquisa e, tampouco seriam julgados pelas respostas do teste. Em seguida, foram entregues os Termos de Esclarecimento e de Consentimento (aprovados pelo Comitê de Ética em Pesquisa da UFTM) para os responsáveis assinarem e, em seguida, as folhas do teste de atitudes.

Para cada parte do teste, houve explicações sobre o funcionamento da atividade e esclarecimento de dúvidas quanto algumas palavras desconhecidas pelos alunos. Houve também orientações para os alunos evitarem conversas paralelas e troca de respostas entre eles. A aplicação do teste durou cerca de quatro aulas (50' cada).

A análise do teste foi feita levando em conta tanto o aspecto quantitativo (quantos alunos responderam de forma igual/semelhante, principalmente na parte 1 e 3 do teste), quanto o aspecto qualitativo (como eles responderam as questões, principalmente na parte 2 do teste, com questões para os alunos responderem por escrito).

Análise do Teste de Atitudes

Nesta seção, apresentaremos a análise de cada parte do teste de atitudes dos 38 alunos que participaram da pesquisa. Para facilitar a análise dos testes e garantir o anonimato, cada aluno participante foi identificado por um código: por turma (T1, T2, etc), e dentro de cada turma, por número (A1, A2, etc).

Análise da primeira parte do teste: perfil social dos alunos

Em relação à primeira parte do teste, ou seja, ao questionário socioeconômico, chegamos aos seguintes resultados:

a) perfil social:

- 67% dos alunos têm entre 12 e 13 anos;

- 60% dos alunos são do sexo feminino;
- 78% dos alunos nasceram em Uberaba, dos quais 50% moram na cidade desde que nasceram e no mesmo bairro da escola (bairro periférico);
- 60% dos alunos moram em uma casa pertencente aos seus responsáveis.
- a média de integrantes da família por casa é de quatro pessoas;
- 60% dos alunos alegaram ter passado dificuldades financeiras.

b) perfil dos responsáveis legais:

- 78% dos alunos têm uma ou duas pessoas da família (responsáveis legais) trabalhando;
- 84% dos responsáveis dos alunos cursaram entre o ensino fundamental e o ensino médio (ou seja, não cursaram nível superior).

c) meios de comunicações e transportes acessíveis (em suas residências):

- 95% das casas dos alunos apresentam pelo menos uma TV;
- 28% dos alunos têm plano de TV por assinatura (a maioria destacou que assiste TV, mas canal aberto);
- 45% dos alunos não têm computador, nem telefone fixo;
- 55% dos alunos não têm automóvel em casa;
- 90% têm acesso à internet via celular.

d) perfil escolar e preocupações dos alunos participantes da pesquisa:

- 55% dos alunos apontarem a saúde como a questão que mais os preocupa, sendo o emprego a segunda questão que mais os preocupa, com 40% das marcações.
- 34% dos alunos já repetiram pelo menos um ano escolar;
- 50% admitiram ter ficado de recuperação ou ter feito alguma prova especial para serem aprovados.
- 50% dos alunos disseram gostar da matéria Português.
- 64% dos alunos afirmaram já terem sofrido algum tipo de discriminação pelo seu jeito de falar na escola.

Cabe destacar dois dados interessantes para nossa pesquisa: apenas metade dos alunos gostarem da disciplina de Língua Portuguesa e 64% deles terem sofrido preconceito linguístico na escola. Por meio da primeira parte do teste, podemos concluir que a maioria dos alunos participantes da pesquisa pertence à classe social mais baixa ou de classe média baixa, e os

responsáveis legais não tem nível superior, levando-nos a inferir que não são falantes da(s) norma(s) culta(s) da região.

Segunda parte do teste (teste escrito): atitudes em relação a tipos de variação

Na segunda parte do teste de atitudes, foi solicitado aos alunos que respondessem, por escrito, três questões. A primeira questão tratava da variação regional e social, a segunda da variação histórica e a terceira da variação geográfica. É importante dizer que apesar de as questões pedirem aos alunos que discorressem sobre o assunto, a maioria dos alunos deu respostas bem curtas, não desenvolvendo bem o seu posicionamento. Houve ainda alunos que deixaram algumas alternativas sem resposta.

Na primeira questão foi-lhes apresentado um trecho transcrito de uma entrevista oral feita com uma mulher que revelava o posicionamento sobre ter filhos:

Eu não, Deus me livre! Eu num posso nem comigo mesmo eu vou tê filho! Tinha um homem que andava atrá de mim, sabe? Disse que queria casá comigo, que a gente ia tê muito fí, pa famía sê grande, mas eu eu quero nada! Pode i simbora procura outra que comigo mesmo não. Quero ganha meu dinheiro, saí, i pa dança, po forró. Oxe, eu gosto é muito de forró! Eu vou cuidá de menino de noite, pa deixa de i: de i po forró. (BAGNO, 2007, p. 208 e 209).

Os alunos deveriam avaliar a fala da mulher (item “a”), mencionar se usam uma forma de falar parecida com a dela (item “b”), qual o estado ou região que consideram ser a entrevistada (item “c”) e ainda de qual classe social que julgavam pertencer a entrevistada (item “d”). Quanto à avaliação da maneira de falar da entrevistada, 45% dos alunos avaliaram bem, apresentando as seguintes respostas:

“Não existe modo errado de falar, apenas certas pessoas trocam as palavras”. (A6T1) / “Maneira informal.” (A5T1) / “Falar correto porque a pessoa entende o que ela fala.” (A8T1) / “Normal.” (A1T2) / “Que ela é de outra cidade.” (A7T2) / “É um jeito diferente, mas é o jeito que ela fala.” (A8T2) / “Que é um modo de falar bonito.” (A1T1)

35% dos alunos avaliaram negativamente a fala da entrevistada, respondendo:

“Ruim” (A3T1 e A7T1) / “Errada” (A2T2 e A4T2) / “Fala diferente o português.” (A3T2) / “Ela podia ser mais educada.” (A6T2)

10% dos alunos consideraram a fala da entrevistada “muito sincera” e o restante não respondeu essa alternativa.

Ainda na primeira questão da segunda parte do teste, 60% dos alunos admitiram usar em sua fala a mesma maneira de falar da entrevistada; 80% dos alunos acreditam que a entrevistada

é de algum estado do Nordeste⁵ e 20% acreditam que a entrevistada é de Minas Gerais. 45% dos alunos consideram a entrevistada integrante da classe média, 35% da classe baixa e os demais não responderam.

Sendo assim, verificamos pelas respostas dos alunos para essa questão que estes ainda consideram “erradas” algumas maneiras de falar e que ainda não têm muita consciência da própria fala, haja vista que mesmo usando na fala algumas das palavras da entrevistada, afirmaram não usar essa forma de falar. Vimos também que acreditam que falantes de classe baixa e média usam uma variedade distante da norma culta e que, para o caso da entrevistada, a maneira de falar indica que se trata de uma pessoa nordestina. Por outro lado, alguns alunos apresentaram uma consciência para o grau de formalidade de uma entrevista oral (transcrita, no caso da questão), não avaliando negativamente a entrevistada. Um dado interessante é que muitos alunos julgaram positivamente a maneira de falar da entrevistada e se identificaram com a variedade usada por ela.

Na segunda questão, foi pedido aos alunos que lessem uma carta escrita pelo poeta Casimiro de Abreu ao pai em 1857:

Rio de Janeiro 11 de Setembro de 1857

Presado Pai

Aqui chegamos sem novidade e tendo felizmente escapado a grande chuva que ainda hoje dura. As cartas todas estão entregues e o Joaquim leva a peça de riscado da loja do primo José e os 12#000, como também uma lamparina ou lampião que me pedi opa comprar. Peço a meu Pai queira recomendar-me ao primo Manoel e ao Senhor Veludo a lançar a sua benção sobre Seu filho

Amante

Casimiro José Marques de Abreu

(BAGNO, M. **Nada na língua é por acaso, por uma pedagogia da variação linguística**. São Paulo: Parábola Editorial, 2007, p.165)

Após o texto da carta, foi perguntado aos alunos o que acharam da linguagem da carta (item “a”) e se atualmente uma carta ao pai seria escrita dessa forma (item “b”), pedindo para os alunos justificarem. Para a primeira alternativa, sobre a avaliação da linguagem da carta, tivemos as seguintes respostas:

“Eu achei a carta muito legal, com linguagem diferente.” (A6T2)/ “Estranha.” (A4T1)/ “Boa.” (A10T2 e A4T2)/ “Muito formal.” (A9T2 e A5T1)/ “Muito boa.” (A3T1)/ “Linguagem mais culta.” (A1T2)/ “Normal.” (A7T2)/ “Sem sentido.” (A6T1)/ “Boa a linguagem e a escrita.” (A8T1)/ “É uma linguagem de antigamente.” (A1T1)/ “Bem escrita.” (A8T2)

Em relação à última alternativa da questão dois, se atualmente uma carta destinada ao pai seria escrita dessa forma, tivemos respostas como:

5 Esse dado é interessante, e deverá ser explorado em pesquisas posteriores.

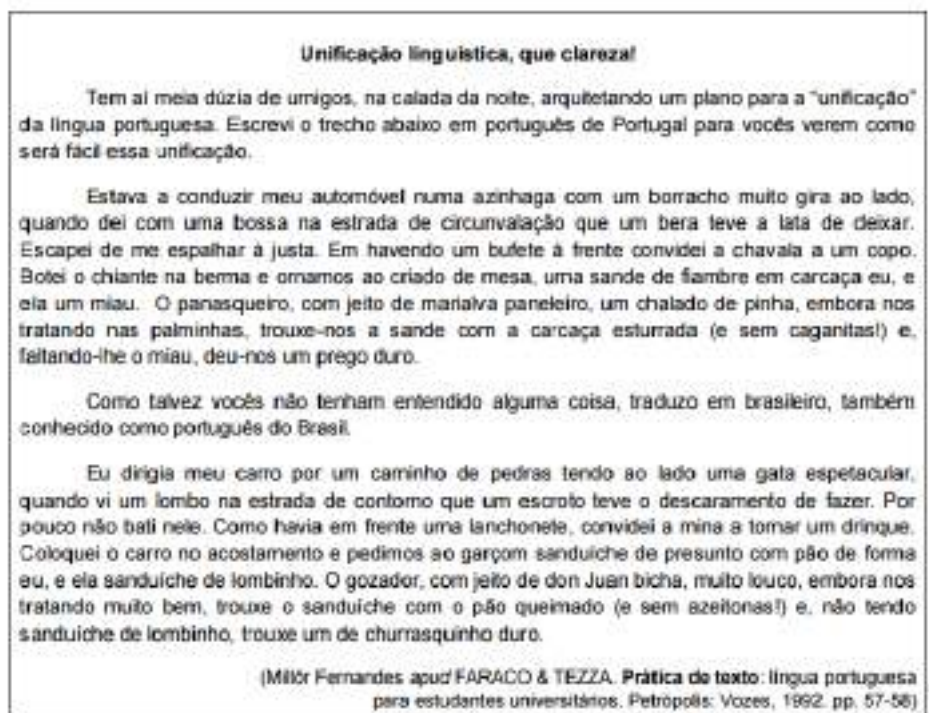
“A carta poderia ter mais informações, mais novidades.” (A6T2)/ “Não seria escrita dessa forma.” (A4T1)/ “Não, porque senão o pai dele não iria entender nada do que ele escreveu.” (A8T1)/ “Não, hoje em dia não falamos com tantas formalidades.” (A5T1)/ “Não, pois primeiro perguntaria se meu pai está bem e contaria se estou bem.” (A7T2)/ “Não, usaria uma linguagem mais simples.” (A1T2)/ “Seria melhor.” (A3T1)/ “Não, mandaria um áudio falando que tava bem.” (A9T2)/ “Sim, porque a linguagem é adequada.” (A4T2)/ “Sim, porque tem que ter educação.” (A2T1)

Como vimos, a maioria dos alunos percebeu que, nos dias atuais, em uma carta escrita ao pai, empregar-se-ia estruturas da língua diferentes da apresentada na questão, ou seja, eles perceberam a variação diacrônica. Apenas 30% afirmaram que atualmente uma carta escrita para o pai usaria a mesma variação da carta de Casimiro de Abreu. Um dos alunos (A9T2) disse que não mandaria carta atualmente, mas sim um áudio, pelas redes sociais, ou seja, levou em consideração o fato de as pessoas não se comunicarem frequentemente por cartas no século XXI.

As respostas dos alunos para a questão dois nos permitem evidenciar que poucos deles apresentaram uma consciência sociolinguística relacionada à variação diacrônica e/ou à mudança linguística, pois não citaram em suas respostas sobre o fato de a língua variar e poder sofrer mudanças ao longo do tempo (chegaram a chamar a carta de “mais formal”, sem considerar questões temporais). A maioria dos alunos somente percebeu que a linguagem usada na carta era “diferente” da atual. Apenas um aluno (A1T1), na alternativa “a”, respondeu ser uma linguagem usada antigamente e na alternativa “b”, apenas um aluno (A9T9) citou o uso dos meios digitais para se comunicar atualmente.

Na última questão, foi mostrado para os alunos um pequeno texto narrativo no português de Portugal, escrito por Millôr Fernandes para tratar da unificação linguística. Foi mostrado aos alunos apenas um trecho do texto integral do autor. Apenas no final da questão é que foi mostrado o texto completo:

Figura 2: Imagem da última questão da segunda parte do teste de atitudes



Fonte: <http://projetoeduc.cecierj.edu.br/eja/recurso-multimedia-professor/portugues/novaeja/mlu02/>

Portugues_do_Brasil_e_portugues_de_Portugal.pdf Acesso em 16/03/2018

Para os itens “a” e “b” do teste, apresentamos aos alunos apenas o segundo parágrafo do texto acima e perguntamos/questionamos se conseguiram entender o texto mesmo escrito em português e por que (item “a”). Em seguida, pedimos que avaliassem o texto (item “b”). Tivemos as seguintes respostas como exemplo:

“Não, porque é muito estranho.” (A4T2)/ “Não, não dá para entender nada.” (A10T2)/ “Não entendi nada, porque são palavras que eu nunca ouvi.” (A9T2)/ “Não pois tem palavras que eu não conheço.” (A1T2)/ “Não, porque ta em outra língua.” (A7T2)/ “Não, porque é um português diferente” (A5T1)/ “Estranho porque eu não entendi.” (A8T1)/ “Estranho, sem sentido.” (A6T1)

“Normal, ele só está em outra língua.” (A7T2)/ “Indescritível.” (A9T2)/ “Ruim.” (A7T1 e A10T2)/ “Estranho.” (A2T2 e A4T2)

Todos os alunos avaliaram negativamente o texto, com exceção de um aluno (A7T2) que considerou o texto normal, mas em outra língua e dos alunos A8T2 e A1T1 que responderam, respectivamente, “Bem escrito” e “Bom”. O estudante participante A1T1, apesar de ter avaliado como “bom” o texto, na letra anterior, afirmou não o ter entendido.

A letra “c” da questão apresentou o texto no português do Brasil e pediu para os alunos escreverem o que concluíram entre os dois textos. Algumas respostas foram:

“Que um dá para entender”. (A10T2)/ “As falas são diferentes.” (A1T1)/ “O texto 1 ficou estranho e o 2 engraçado.” (A2T2)/ “Um fala uma língua e o outro fala outra língua.” (A5T2)/ “O segundo fica mais na linguagem que nois entende.” (A9T2)/ “ O segundo texto tem palavras que eu conheço, portanto eu entendi.” (A1T2)/ “O texto 2 está melhor escrito que texto 1.” (A6T1)/ “Um é escrito com o português brasileiro e o outro não.” (A5T1)/ “O segundo texto está melhor que o primeiro.” (A8T1)

Como mostrado, os alunos consideraram o texto em português brasileiro melhor, mais compreensível, mas não discorreram muito sobre a variação linguística e o porquê dos dois textos, mesmo estando em uma mesma língua, serem muito diferentes. Quatro alunos não responderam a letra “c”.

As respostas da questão três, última questão da segunda parte do teste, mesmo utilizando um texto ficcional, sem mostrar um uso real da língua portuguesa (tanto em Portugal, quanto no Brasil), conseguimos observar que a maioria dos alunos avalia como “ruim” e “estranho” a variedade do outro, principalmente, a que não conhecem. A variedade familiar a eles é avaliada como “melhor”.

De modo geral, as três questões da parte escrita do teste mostraram que os alunos têm muita dificuldade em reconhecer a variação linguística e ter um posicionamento sensível à variedade do outro. Isso prova o quanto o assunto – variação linguística - não é muito abordado em sala, no livro didático, enfim, na escola. Nenhum aluno utilizou palavras como “variação”, “variedade” para explicar a maneira de falar mostrada nos textos. Além disso, percebemos o quanto o preconceito linguístico ainda está presente no ambiente escolar, afinal muitos alunos avaliaram negativamente as variedades mostradas nos textos.

Nesse sentido, a segunda parte do teste evidenciou o quanto é necessário um trabalho sociolinguístico em sala de aula. A percepção dos alunos em relação às variedades linguísticas é ainda muito sutil. É preciso um trabalho que mostre aos alunos as diversas variedades da nossa língua e o quanto a diferença presente nessas variedades é positiva. É imprescindível mostrar-lhes que a variação é uma característica inerente a todas as línguas. Felizmente, percebemos pelas respostas de certos alunos que alguns já apresentam uma consciência sociolinguística e consideram “normais” as variedades da língua e que não julgam tais variedades apenas por serem diferentes da norma ensinada pela escola. Verificamos nessa parte do teste que os alunos apresentam uma reflexão sobre a língua, porém ainda muito precária. E ainda consideram determinadas maneiras de falar esquisitas, ruins e estranhas.

Terceira parte do teste: avaliação das gravações

Na terceira parte do teste de atitudes, os alunos escutaram quatro gravações de falantes diferentes:

Falante 1: masculino, 55 anos, fundamental incompleto

Falante 2: feminino, 80 anos, ensino fundamental incompleto

Falante 3: masculino, 25 anos, superior completo

Falante 4: feminino, 29 anos, ensino superior completo

Os alunos não sabiam nenhuma característica desses falantes e deveriam avaliá-los, pela maneira de falar, quanto aos seguintes adjetivos (que representam diferentes campos semânticos): Inteligente, Honesto, Competente, Simpático, Rico e Boa pessoa. Os alunos receberam a seguinte instrução:

Quadro 1: fragmento da terceira parte do teste de atitudes

Quando ouvimos a voz de uma pessoa, sem vê-la, formamos impressões sobre ela. É o que provavelmente acontecerá com você quando ouvir os trechos gravados que lhe serão apresentados agora. São quatro gravações, cada uma delas trazendo a voz de pessoas diferentes. Depois de ouvir cada um desses pequenos trechos, indique a impressão que lhe causou, a partir das questões abaixo.

A sua resposta deve ser dada da seguinte maneira: atribua uma nota variando de 7 a 1, em relação às qualidades expressas pelos adjetivos que estão relacionados. Você tem seis adjetivos, cada um deles seguido de sete traços. Cada traço corresponde a uma nota.

Quanto mais próxima do adjetivo, maior será a nota. Assim, se você atribuir a um dos falantes a nota mais alta (7), em relação ao adjetivo inteligente, você deve colocá-la no traço próximo a esse adjetivo; se você atribuir a nota imediatamente inferior (6), deve colocá-la no segundo traço, e assim sucessivamente, até a menor (1), que deverá ser colocada, se for o caso, no último traço.

Você vai ouvir uma gravação de cada vez. Depois de cada uma delas, vamos lhe dar um tempo para você colocar as notas nos traços:

FALANTE 1

Inteligente ____: ____: ____: ____: ____: ____: ____

Honesto ____: ____: ____: ____: ____: ____: ____

Competente ____: ____: ____: ____: ____: ____: ____

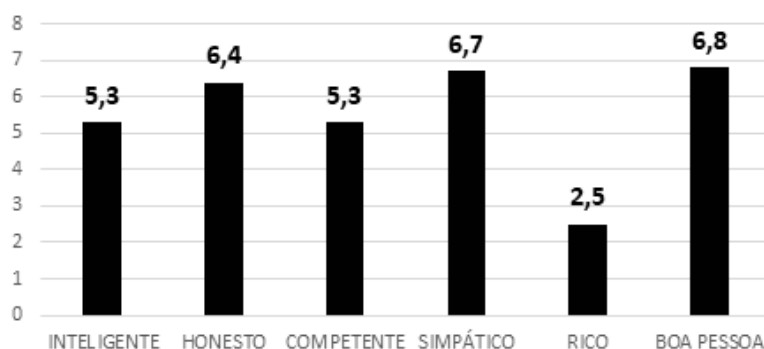
Simpático ____: ____: ____: ____: ____: ____: ____

Rico ____: ____: ____: ____: ____: ____: ____

Boa pessoa ____: ____: ____: ____: ____: ____: ____

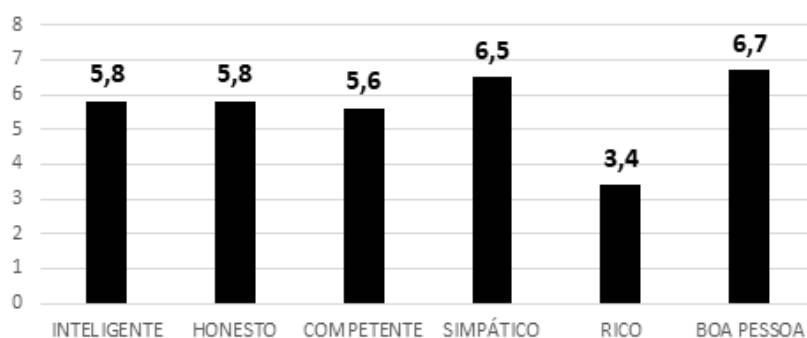
Fonte: extraído do teste de atitudes elaborado e aplicado pelas autoras

A partir das respostas obtidas, realizamos análises quanti-qualitativas com auxílio do Excel. Em relação ao falante 1, a seguir mostramos o gráfico 1 com a média das notas que recebeu dos alunos para os seis adjetivos:

Gráfico 1: Média das notas do Falante 1 (Terceira parte do Teste de Atitudes Linguísticas)

Fonte: as autoras

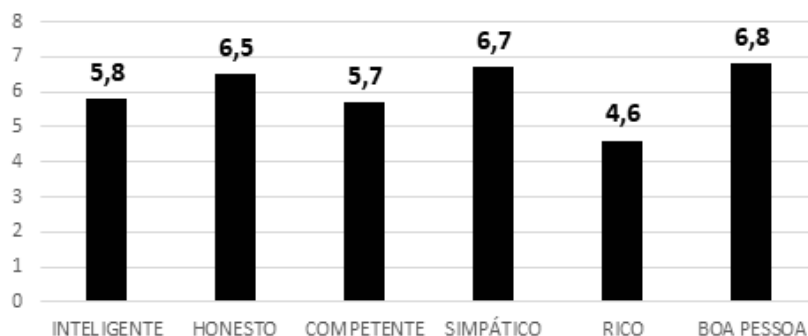
De acordo com o gráfico 1, o Falante 1, do sexo masculino, de 55 anos e cuja escolaridade é nível fundamental incompleto, recebeu nota mais baixa para o adjetivo “rico” (2,5 de 7,0 pontos). As outras duas notas mais baixas foram para “inteligente” (5,3) e “competente” (5,3). A variedade usada pelo Falante 1 (transcrição da gravação do falante está no Anexo II) não é uma variedade prestigiada. Além disso, a entrevista da qual participou permitiu o uso de uma linguagem informal e espontânea por tratar de memórias da infância. Apesar dos alunos avaliarem, pela maneira de falar, o falante como pouco rico, não muito “inteligente” e nem tanto “competente”, os alunos deram notas mais altas para os adjetivos “honesto”, “simpático” e “boa pessoa”, indicando que, para eles, o uso de uma variedade “não-culta” diz sobre a condição social da pessoa e sua inteligência, mas não interfere em sua simpatia, honestidade e caráter (“boa pessoa”). Para o Falante 2, temos o seguinte gráfico de notas para os adjetivos:

Gráfico 2: Média das notas do Falante 2 (Terceira parte do Teste de Atitudes Linguísticas)

Fonte: as autoras

Para o Falante 2, de sexo feminino, de 80 anos, cuja escolaridade é ensino fundamental incompleto, de variedade não prestigiada (apesar do monitoramento da fala no momento da entrevista, principalmente no início) também foi atribuída nota mais baixa para o adjetivo “rico”. As notas mais altas continuaram, assim como para o Falante 1, para os adjetivos “simpático” e “boa pessoa”. O Falante 2 foi considerado mais inteligente que o Falante 1, porém menos honesto. Os dados ainda indicam que, para os alunos, a fala de uma pessoa diz sobre a condição social/econômica e sobre sua inteligência, mas não diz negativamente sobre outras qualidades. A seguir a média das notas do Falante 3:

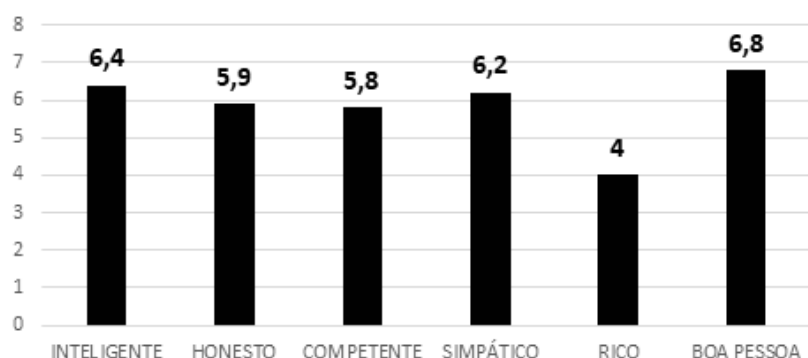
Gráfico 3: Média das notas do Falante 3 (Terceira parte do Teste de Atitudes Linguísticas)



Fonte: as autoras

O Falante 3, diferentemente dos anteriores, era mais jovem e com o nível de escolaridade maior. É do sexo masculino, 25 anos e tem ensino superior completo. Usou também uma variedade não prestigiada, com algumas gírias. Os alunos também atribuíram ao Falante 3, como aponta o gráfico 3, uma nota baixa para o adjetivo “rico” (4.6), porém mais alta do que as notas do Falante 1 e do Falante 2. A segunda nota mais baixa foi para o adjetivo “competente” (5,7) e em seguida para o adjetivo “inteligente”. As notas mais altas foram para os adjetivos “honesto”(o mais honesto entre os falantes), “simpático” e “boa pessoa”. Em relação ao Falante 4, temos o gráfico 4:

Gráfico 4: Média das notas do Falante 4 (Terceira parte do Teste de Atitudes Linguísticas)



Fonte: as autoras

O Falante 4, de sexo feminino, de 29 anos e também com ensino superior completo, foi o falante com maior monitoramento da fala ao contar sobre as memórias de sua infância. Podemos dizer que sua variedade está mais próxima da variedade considerada culta da região. Foi o falante que ganhou a nota mais alta para o adjetivo “inteligente” (6,4), mas sua nota para o adjetivo “rico” (4) foi mais baixa que a do Falante 3. Também foi considerado “honesto”, “competente”, “simpático” e “boa pessoa”. É interessante observarmos que, de acordo com a nota dos alunos, o fato de o Falante 4 usar uma variedade mais próxima da culta não o faz ser considerado mais rico do que os outros, afinal, o Falante 3 obteve a nota mais alta para o adjetivo “rico”, mesmo utilizando uma variedade “não-culta” e um estilo mais informal que o Falante 4.

Como fica evidente na análise dos gráficos de 1 a 4, a maioria dos alunos consideraram os falantes “inteligentes, honestos, competentes, simpáticos e boas pessoas”, atribuindo notas mais altas para esses adjetivos. Sendo os adjetivos “simpático” e “boa pessoa” os que ganharam maiores notas entre os quatro falantes. O único adjetivo que teve nota mais baixa em relação aos demais adjetivos foi o “rico”. Isso nos mostrou que os alunos consideram a forma de falar de uma pessoa muito ligada com sua condição social/econômica, mas que a condição social não interfere no caráter do falante.

Devemos considerar também como variante para esses dados, não só a maneira de falar (a variedade) do falante, mas também o conteúdo de cada gravação. Dizemos isso porque ao aplicarmos a última parte do teste, alguns alunos comentaram espontaneamente sobre a realidade de vida dos falantes na infância. Não considerando que, talvez, na fase adulta, os falantes possam ter mudado sua condição social. Afinal, alguns falantes disseram ter passado uma infância pobre, sem muitos recursos. Apesar dessa questão, consideramos que, segundo a avaliação dos alunos, a variedade de uma pessoa diz muito sobre sua condição social/econômica, mas não sobre outras qualidades/características dela.

Resultados da Aplicação

Apontamos os resultados gerais obtidos no teste de atitudes linguísticas aplicado nas turmas de 7º Ano de uma escola pública de Uberaba (MG). A aplicação do teste de atitudes nos permitiu comprovar que:

- (i) muitos alunos ainda consideram determinadas maneiras de falar “erradas”, “estranhas”;
- (ii) há ainda alunos que não veem/percebem a diferença linguística, a heterogeneidade da língua, como algo positivo e enriquecedor para pesquisas linguísticas;
- (iii) o preconceito linguístico ainda está presente nas escolas (e na sociedade também), haja vista que alguns alunos afirmaram já ter sofrido preconceito linguístico (no teste) e ainda demonstraram preconceito (no teste);
- (iv) muitos alunos desconheciam, apesar de cursarem há sete anos aulas de Língua Portuguesa na escola, o que é variação linguística;
- (v) os alunos acreditam que a maneira de falar de alguém não está relacionada com seu caráter;
- (vi) os alunos julgam a condição social de uma pessoa de acordo com sua maneira de falar.

Ponderações Finais

Em nossa pesquisa, procuramos mostrar os principais pontos dos trabalhos e estudos lidos a respeito do tema da Sociolinguística Educacional e das atitudes linguísticas. Mostramos a relevância/necessidade/importância de se trabalhar com a variação linguística na escola, principalmente nas aulas de Língua Portuguesa - haja vista as inúmeras contribuições que o estudo das variações linguísticas pode dar aos alunos, como preconizaram os documentos oficiais do governo há anos – e ainda descrevemos a aplicação do teste de atitudes e seus resultados.

Esperamos ter evidenciado a relevância e a atualidade do tema da pesquisa (proposta de ensino de Língua Portuguesa mais sensível às variações linguísticas) e buscamos colaborar com pesquisas já feitas e também com as que surgirão. Afinal, levar conhecimentos da Sociolinguística Educacional às escolas e conhecer as atitudes/avaliações linguísticas de alunos são de grande importância para fazermos o ensino de língua materna progredir e formar pessoas competentes no uso de sua própria língua.

Referências

AGUILERA, V. A. Crenças e atitudes linguísticas: o que dizem os falantes das capitais brasileiras. In: *Estudos Linguísticos*, São Paulo, 37 (2): 105-112 maio-agosto, 2008.

BAGNO, M. *Nada na língua é por acaso: por uma pedagogia da variação linguística*. São Paulo: Parábola Editorial, 2007.

BARCELOS, A. M. F; ABRAHÃO, M. H. V. *Crenças e ensino de línguas: foco no professor, no aluno e na formação de professores*. Campinas: Pontes, 2006.

BARBOSA, J. B.; CUBA, D. L. 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, 2015.

BORTONI-RICARDO, S. M. *Educação em língua materna: a Sociolinguística na sala de aula*. São Paulo: Parábola Editorial, 2004.

_____. *Nós chegemos na escola, e agora?* Sociolinguística e educação. São Paulo: Parábola Editorial, 2005.

BOTASSINI, J. O. M. A Importância dos Estudos de Crenças e Atitudes para a Sociolinguística. *SIGNUM: Estud. Ling.*, Londrina, n. 18/1, p. 102-131, jun. 2015.

BRASIL. Ministério da Educação. *Parâmetros Curriculares Nacionais: Língua Portuguesa*. Brasília: MEC, 1997.

BRASIL. Ministério da Educação. *Parâmetros Curriculares Nacionais: Língua Portuguesa*. Brasília: MEC, 1998.

BRASIL. Base Nacional Comum Curricular (BNCC). Educação é a Base. Brasília, MEC/ CONSED/UNDIME, 2018.

CORBARI, C. C. Atitudes linguísticas: um estudo nas localidades paranaenses de Irati e Santo Antônio do Sudoeste. Tese apresentada ao Curso de Doutorado em Letras, do Programa de Pós-Graduação em Letras e Linguística da Universidade Federal da Bahia – UFBA, Salvador 2013.

CYRANKA, L. F. M. Atitudes linguísticas de alunos de escolas públicas de Juiz de Fora-MG. 2007. Tese (Doutorado em Estudos Linguísticos). Universidade Federal Fluminense, Niterói.

CYRANKA, L. F. M. Sociolinguística aplicada à educação. In: MOLLICA, M. C.; FERRAREZI R, C. *Sociolinguística, Sociolinguísticas*. São Paulo: Editora Contexto, 2016.

CYRANKA, L. F. A pedagogia da variação linguística é possível? In: ZILLES, A M. S.; FARACO, C. A. *Pedagogia da variação linguística: língua, diversidade e ensino*. São Paulo: Parábola Editorial, 2015.

CUBA, D. L. Uma proposta de ensino de língua portuguesa a partir da identificação de atitudes linguísticas de alunos do 7º ano de uma escola pública de Uberaba-MG. Mestrado (Programa de Pós-Graduação em Mestrado Profissional em Letras – PROFLETRAS). 2019. Universidade Federal do Triângulo Mineiro. Uberaba, MG.

ECKERT, P. Three Waves of Variation Study: the emergence of meaning in the study of Variation. *Annual Review*. 41, 2012.

FARACO, C. A. *Norma culta brasileira: desatando alguns nós*. São Paulo: Parábola Editorial, 2008.

_____. Por uma pedagogia da variação linguística. In: CORREA, D. A. (org). *A relevância social da linguística: linguagem, teoria e ensino*. São Paulo: Parábola Editorial, 2007.

_____. *Por uma pedagogia da variação linguística*. 2011. Disponível em:

<https://variacaolingustica.files.wordpress.com/2011/06/faraco_por_uma_pedagogia_da_variacao_linguistica1.pdf>. Acesso em: 12/07/2017.

GONZÁLEZ, C. A. Variação linguística em livros de português para o EM. In: ZILLES, A. M. S; FARACO, C. A. *Pedagogia da variação linguística: Língua, diversidade e ensino*. São Paulo: Parábola Editorial, 2015.

IDEB. Índice de Desenvolvimento da Educação Básica, 2017. Disponível em: <http://ideb.inep.gov.br/Site/> Acesso em 20/10/2018.

LAMBERT, W. E. et al. Evaluation reactions to spoken languages. *Journal of Abnormal Social Psychology*, n. 60, p. 44-51, 1960.

LIMA, R. J. Variação linguística e os livros didáticos de português. In: MARTINS, M. A.; VIEIRA, S. R.; TAVARES, M. A. *Ensino de português e sociolinguística*. São Paulo: Contexto, 2016.

MARINE, T. de C.; BERTUCCI BARBOSA, J. Em busca de um ensino sociolinguístico de Língua Portuguesa no Brasil. *Signum: Estudos da Linguagem*, [S.l.], v. 19, n. 1, p. 185-215, jan. 2017. ISSN 2237-4876. Disponível em: <<https://www.uel.br/revistas/uel/index.php/signum/article/view/23161>>. Acesso em: 22 jun. 2021.

MARQUES, T. M; BARONAS, J. E. A. Crenças e atitudes linguísticas em sala de aula. *Linguasagem*, São Carlos, v. 24 (1): 2015.

MENDES, R. B. A terceira onda da Sociolinguística. In: FIORIN, J. L.(org). *Novos caminhos da linguística*. São Paulo: Contexto, 2017.

MORENO FERNÁNDEZ, F. *Principios de Sociolinguística y Sociología Del Lenguaje*. Barcelona: Ariel, 1998.

ZILLES, A. M. S; FARACO, C. A. (org). *Pedagogia da variação linguística: língua, diversidade e ensino*. São Paulo: Parábola Editorial, 2015.



SUBJECT EXPRESSION IN SPOKEN VARIETIES OF PORTUGUESE

A EXPRESSÃO DO SUJEITO EM VARIEDADES FALADAS DO PORTUGUÊS

*Aline Maria Bazenga*¹

ABSTRACT

Research on spoken varieties of Portuguese allows for the observation of diverse socio-historical and cultural profiles in communities where the language is spoken, and – in the field of linguistic systems – of the occurrence of variable morphosyntactic phenomena. In this article, particular attention is given to the specific phenomenon of variable subject expression. The study is based on a theoretical framework that brings together contributions from discrete theoretical models of Linguistic Variation (Generative Grammar vs. Linguistic Typology). Taken together, the contributions of these models allow for a characterisation of the complexity of the variable expression of the subject in varieties of Portuguese. A third perspective is also added to these two: one in which linguistic variation is related not only to linguistic factors but also to social variation. Within the framework of the model proposed by variationist sociolinguistics, the article seeks to draw a comparison between varieties, to trace the possibilities for a generalisation of null subject and pronominal subject production, to argue for the visualisation of such production along a *continuum*, and to contribute to broadening our understanding of the linguistic phenomenon in question. The results of this comparison of fifteen published works, which were selected due to their common deployment of variationist / quantitative analysis based on *corpora* of Portuguese varieties, points to: i) a highly complex situation in the sphere of Portuguese when non-standard varieties are taken into account, similar to what has been observed in other romance languages such as Italian; ii) the need for a common methodological framework in order to produce a greater quantity of comparable linguistic data; and iii) the need for further theoretical reflection, given the scant attention paid to social or extralinguistic factors.

KEYWORDS: Variable subject expression; Spoken varieties of Portuguese; Variationist and sociolinguistic approaches; Linguistic variation; Null-Subject Languages.

¹ Associate Professor in the Department of Languages, Literatures, and Cultures, Faculty of Arts and Humanities, University of Madeira. E-mail: alinebazenga@staff.uma.pt.

RESUMO

A investigação sobre as variedades faladas do português permite-nos observar uma diversidade de perfis histórico-sociais, culturais das suas comunidades, como também, no plano dos sistemas linguísticos, a ocorrência de fenómenos morfossintáticos variáveis. Neste artigo, uma atenção particular será dada a um destes fenómenos: a expressão variável do sujeito. Parte-se de um enquadramento teórico que reúne contributos oriundos de modelos teóricos sobre a Variação Linguística distintos (Gramática Generativa vs Tipologia Linguística), mas cujos contributos permitem configurar o objeto de estudo – a expressão variável do sujeito em variedades do português – sua complexidade. A estes duas perspetivas associa-se uma terceira, na qual a variação linguística está correlacionada não apenas com fatores linguísticos, mas também com a variação social. Pretende-se, no quadro do modelo proposto pela sociolinguística variacionista, estabelecer uma comparação entre variedades, traçar possibilidades de generalização de produção de sujeitos nulos e de sujeitos pronominais, propondo o seu alinhamento em *continua*, e contribuir para o conhecimento mais alargado do fenómeno linguístico em questão. Os resultados, com base em cerca de quinze trabalhos selecionados por terem em comum o facto de recorrerem a análises variacionistas/quantitativas, a partir de *corpora* oral de variedades do português, apontam para (i) uma situação de significativa complexidade na esfera do português, à imagem do observado noutras línguas românicas, como o italiano, quando se têm em conta variedades não padrão, (ii) para a necessidade de um protocolo metodológico em comum, por forma a produzir maior quantidade de dados linguísticos comparáveis e (iii) a necessidade de reflexão teórica, face à pouca relevância observada de fatores sociais ou extralinguísticos.

PALAVRAS-CHAVE: Expressão variável do sujeito; Variedades faladas do português; Abordagens variacionistas e sociolinguísticas; Variação linguística; Línguas de Sujeito Nulo.

Introduction

The diversity within varieties of Portuguese has been widely studied at a lexical and phonetic level and also, although more recently, in the field of syntax, not only from a formal linguistic perspective but also from the starting point of non-formal, functionalist, parametric, and variationist approaches based on empirical data.

The present study seeks to draw a large-scale comparison between the different forms of expressing the subject in varieties of Portuguese, based on the results of studies that have taken non-formal approaches. The Portuguese language data compiled here include not only social and geographical varieties of L1 Portuguese, but also L2 varieties. From the outset, this comparison is a highly complex task (Miestamo, Sinnemäki & Karlsson, 2008; Sampson, Gil & Trudgill, 2009). The variable in question is of a complex nature, as has been frequently observed in the literature on formal and generative linguistics; also, the variety of methodologies adopted in the treatment of empirical data in non-formal studies also complicates this exercise of analysis and comparison of research findings. In short, and as Arends (2001: 180) observes, “more is more complex” – and that is certainly the case in this study.

This article aims to contribute to a better understanding of the variable phenomenon in question across a set of Portuguese varieties for which research data has already been published. The research is divided into two sections, followed by a conclusion. The first section commences with a presentation of the *subject expression* variable within the broader context of interlinguistic variations and takes account of the patterns observed in the Linguistic Typology,

and the most recent discussion of this as a *null subject parameter* within the scope of Generative Grammar. This first section also discusses the variability of subject expression across the group of romance languages within which Portuguese is situated. The second part of the article makes a comparative examination of how the linguistic variable in question functions in different varieties of Portuguese. The analysis proposed here seeks to foreground the convergences and divergences between varieties, insofar as this exercise is possible.

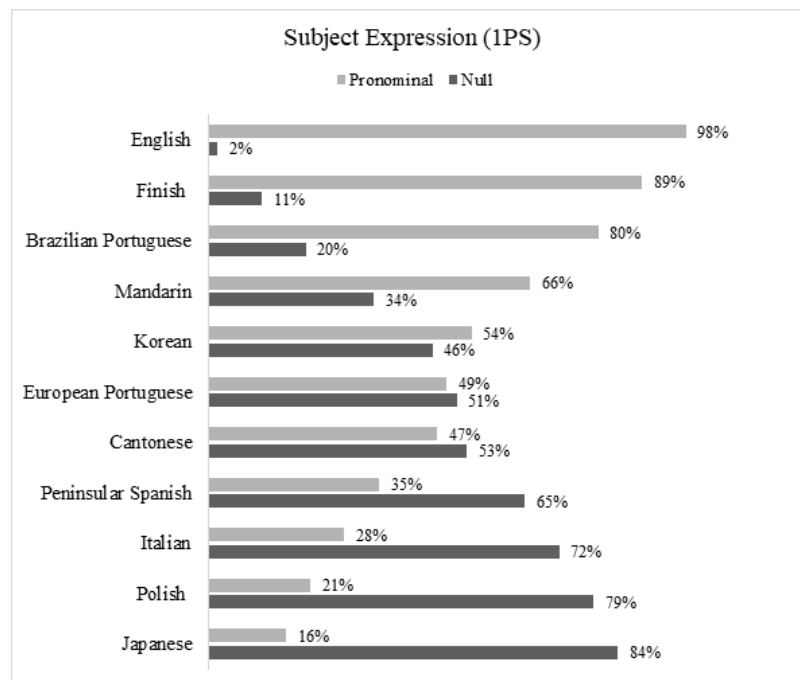
Subject expression in natural languages: a variable linguistic phenomenon

It is widely known that in a language such as Spanish, for example, the subject of a sentence may be expressed overtly as a noun (1a.) or pronoun (1b.); or it may be omitted and not overt, or null (1c.). English, by contrast, allows only the overt expression of the subject, as exemplified in (2).

- (1) a. *El niño* estaba estudiando
b. *Yo* estoy cantando
c. \emptyset estoy cantando
- (2) a. *The child* was studying.
b. *I* am singing.

Graph 1 (below) observes this variable phenomenon across very diverse languages such as Japanese, Italian, Finnish, and English.

Graph 1. Null subject variation in 1PS in typologically different languages.



Source: Torres Cacoullos & Travis (2019), whose work is based on the following studies: for Japanese 84% (Lee & Yonezawa, 2008: 738, N=1571), Polish 79% (Chociey, 2011: 52, N=536), Italian 72% (Nagy p.c. cf. Nagy et al. 2011, N=224), Peninsular Spanish 65% (Posio, 2013: 269, N=787), Cantonese 53% (Nagy p.c. cf. Nagy et al. 2011, N=362), European Portuguese 51% (Posio, 2013: 269, N=704), Korean 46% (Oh, 2007: 466, N=433), Mandarin 34% (Jia & Bayley, 2002: 13, N=393), Finnish 11% (Helasvuo, 2014: 68, N=1793), English 2% (Torres Cacoullos & Travis, 2014: 22, N=6,600 (estimated)), and Duarte (2020), for two varieties of Portuguese (European Portuguese, 51% and Brazilian Portuguese, 20%).

Within the methodological framework employed in the World Atlas of Language Structures (WALS)² project (Dryer & Haspelmath, 2013), Dryer (2013) establishes six types of languages in accordance with the properties that they present in terms of subject expression in declarative sentences (Table 1).

Table 1. Subject expression and language typology (Dryer, 2013).

	Features / Properties	N	Language
I	Obligatory pronouns in the position of subject	82	English, French
II	Subjects expressed by verb affix	437	Turkish, Spanish, Italian,
III	Subjects expressed by clitics on variable host	32	Chemehuevi
IV	Subject pronouns that occur in a different position from that of the subject	67	Longgu
V	Pronouns in the subject position but often omitted.	61	Japanese
VI	More than one trend without a dominant pattern	32	Finnish
	Total	711	

Source: Dryer (2013)

Type I languages, such as English in example (2) above, and French, are characterised by the realisation of the pronoun in the position of the subject. Type II, which includes European Portuguese, is the most frequently occurring paradigm. More than half of the 711 languages examined by Dryer fall into this category and have as a common feature the fact that the subject marker can be realised through verb affixes. The majority of romance languages are distributed across these first two types; most of them follow the dominant trend of this typology and fall into type II, but French is a type I language (Table 2).

Table 2. Typological classification of romance languages based on Dryer (2013).

Romance languages	Pronoun	Typological classification (Dryer, 2013)
European Portuguese	<i>(ele) come uma maçã</i>	Type II
European Spanish	<i>(el) come una manzana</i>	
Italian	<i>(lui) mangia una mela</i>	
Romanian	<i>(el) mananca un mar</i>	
Catalan	<i>(es) menja una poma</i>	
French	<i>*(il) mange une pomme</i>	Type I
Romansh	<i>*(Jeu) magliel in meil.</i>	

Source: Kaiser & Hack, 2010: 83

According to Kaiser & Hack (2010: 84), within the romance branch a third category exists in addition to these two types that should form part of type VI (Table 1), and for which the obligatory nature of the subject pronoun is conditioned by its morphological category of PERSON.

² WALS project: <https://wals.info/>

Toutes les études sur les dialectes septentrionaux de l'Italie s'accordent sur le fait que l'emploi du pronom sujet (clitique) dépend de la personne grammaticale, c.-à-d. qu'il est obligatoire ou quasi-obligatoire avec certaines personnes grammaticales, tandis que d'autres personnes présentent une variation libre par rapport à la présence du pronom (Heap 2000). La même conclusion peut être tirée par rapport à certaines variétés de l'occitan (SIBILLE 2007) ou du francoprovençal (DE CROUSAZ e SHLONSKY 2003; DIÉMOZ 2007; HINZELIN and KAISER 2008)³.

For these authors, Romansh should not be included in Type I because it presents grammatical properties that distinguish it from French as much as from the other romance languages in Table 1. Unlike French, Romansh is a language with a rich verbal morphology. On the other hand, although the pronominal realisation of the subject is obligatory, it does not obey the same conditions as in French, for its omission is possible in contexts where it is post-verbal and in agreement with the grammatical PERSON. Such omission is very frequent in the second person, which makes Romansh an unusual language within the romance family, because it can also be included within type VI of Dryer's (2013) Typology.

C'est l'omission du pronom sujet de la deuxième personne du singulier qui constitue la différence cruciale entre le romanche et les variétés romanes à *pro-drop* partiel: alors que l'omission de ce pronom est extrêmement fréquente en romanche, sa réalisation est obligatoire dans les autres variétés romanes – et encore plus obligatoire que celle de tout autre pronom sujet (KAISER & HACK, 2010: 90)⁴

This variable phenomenon has been widely studied using the theoretical model of Principles and Parameters of generative grammar, and it is within this framework that it has been interpreted as a Null Subject Parameter (RIZZI, 1982: 117-184; JAEGGLI & SAFIR, 1989: 1-44). From this perspective, a language such as Italian has the property [+Null Subject], with the indicators of a *pro-drop* language; it permits the omission of the subject therefore it is a Null-Subject Language (NSL). French is distinct from the parameter of Italian in that it is a non-Null-Subject Language (non-NSL) or [-Null Subject]. In contrast with Italian, French does not allow the omission of the subject. Nevertheless, and as Rizzi (1986) points out, it is not possible to account for interlinguistic variation only by establishing an opposition between

3 All studies of northern Italian dialects agree that the use of the subject pronoun (clitic) depends on the grammatical person, i.e., that it is obligatory or quasi-obligatory with certain grammatical persons, while other persons present free variation with respect to the presence of the pronoun (Heap, 2000: 66). The same conclusion can be drawn with respect to certain varieties of Occitan (Sibille, 2007) or Franco-provençal (De Crousaz & Shlonsky, 2003; Diémoz, 2007; Hinzelin & Kaiser, 2008).

4 It is the omission of the second-person singular subject pronoun that constitutes the crucial difference between Romansh and the partial *pro-drop* Romance varieties: while the omission of this pronoun is extremely frequent in Romansh, its realisation is obligatory in other romance varieties – and even more obligatory than that of any other subject pronoun (Kaiser & Hack, 2010: 90).

the obligatory vs. possible expression of the subject. As outlined above, other grammatical properties such as the grammatical category of person/number in the verb and/or the inversion of the position of the subject can have an influence, for example impeding the categorisation of a language as a NSL. In addition to needing to observe the verbal morphology, whether rich or not, of a given language, it is also necessary to take into account the different types of subject and their correlation with the alternation between null subject and pronominal subject. As such, it is necessary to distinguish between different types of subject:

- (3) a. **null-argument, expletive subject with no referential content** (*il semble que le temps va changer*)
- b. **argument subject, but one that is arbitrary or has an indefinite interpretation** (*On frappe à la porte*)
- c. **argument subject with a definite reference** (*j'ai lu tout le livre*).

While the pronominal subject is obligatory in all contexts in French (cf. types of subjects) as illustrated in (1), the same cannot be said of standard European Portuguese. In this variety of Portuguese, the realisation of the null subject is obligatory in most contexts / for most types of subjects, although it is optional for argument subjects with a defined reference. These three types of subject may be *pro* or null, depending on the language, and they are characterised by a range of properties (Table 3).

Table 3. Types of subjects

Interpretation (types) of <i>pro</i>	PERSON	NUMBER
<i>pro</i> expletive	+	+
<i>pro</i> quasi-argument	-	+
<i>pro</i> expletive	-	-

Source: Sheehan (2016: 333)

Holmberg (2010) discusses a set of properties that interfere in the demarcation of the null subject parameter and proposes five types of languages (Table 4): in addition to the consistent or classical NSLs (Portuguese, Italian, Spanish) and the non-NSLs such as French and English that are described above, it is also important to consider partial NSLs (such as Finnish), expletive NSLs (German), and radical NSLs (Chinese, Japanese).

Table 4. Types of subjects and Homberg's (2010) typology

Types of Subjects	Italian, European Portuguese, Greek	English, French	Finnish	German, Dutch	Chinese, Japanese
null-argument subject, expletive, with no referential content	Not expressed	Expressed	Not expressed	Not expressed	Not expressed
quasi-argument subject, arbitrary or of indeterminate interpretation	Not expressed	Expressed	Variable	Expressed	Not expressed
Argument subject with definite reference	Variable	Expressed	Variable	Expressed	Not expressed
Holmberg's (2010) typology	I Consistent null-subject languages	II Non-null-subject languages	III Partial null-subject languages	IV Expletive null-subject languages	V Radical null-subject languages

Sources: Lobo (2016); Holmberg (2010); Figueiredo Silva (2017).

Another question that has attracted the attention of researchers is related to the different interpretations that can convey null subjects and overt subjects in NSLs (MONTALBETTI, 1984). These interpretations are related to pragmatic-discursive properties (CARDINALETTI & STARKE, 1999) and gave rise to the *Antecedent Position Hypothesis* (CARMINATI, 2002). In line with this hypothesis, Madeira, Fiéis & Teixeira (2021) observe that in European Portuguese, null subjects tend to refer to antecedents in the position of the subject, as in (4a.), while overt subjects tend to recover antecedents in other syntactic positions or extralinguistic referents, as in (4b.).

- (4) a. A menina_i beijou a avó_j quando *pro*_i⁵ chegou a casa
b. A menina_i beijou a avó_j quando *ela*_j/_k chegou a casa

Madeira, Fiéis & Teixeira (2021: 156)

5 *Pro* is a term used within the field of generative grammar, and especially in *Lectures on Government and Binding* (Chomsky, 1981), that denotes an empty category to account for phrasal elements that are not uttered (that are not produced phonetically) but have an effect on syntax. In this case, *pro* takes the subject position and receives the nominative case. *Pro* also forms part of a grammatical parameter – the Null Subject Parameter – of this theory, which seeks, amongst other things, to account for commonalities and divergences across all languages. From this perspective, in null subject languages (or *pro-drop* languages) a positive value is attributed to the parameter; while the languages in which the parameter is given a negative value are those which do not allow the omission of the subject, in other words, obligatory pronominal subject languages.

These authors also point out that the interpretative tendency of pronominal subjects can be cancelled by pragmatic factors, as **illustrated** in (5), whereby the null subject refers not to the constituent subject, but to an extra-phrasal antecedent:

- (5) O João_i não viu o despertador. O Rui_j pensa que **pro**_i vai chegar atrasado.

Lobo (2013: 2323)

From the perspective of the formal model of generative grammar, the theoretical treatment of the Null Subject Parameter presents at least two challenges, which are observed by D'Alessandro (2014) and discussed further by Figueiredo Silva (2017: 194): on the one hand, there is a need to respond to empirical counterexamples, and in particular those which are based on dialectal data from European NSLs (such as varieties of Italian), and/or non-European varieties of Spanish (CAMACHO, 2013, for Dominican Spanish, among others), and/or other non-European NSLs (GILLIGAN, 1987); on the other, there is a need to resolve the problem presented by theoretical counter-evidence that jeopardises “the very existence of an empty category of the type of *pro*” and “the difficulty of attributing a clear theoretical status” to the micro-parameter of inflectional “richness”. Even so, D'Alessandro underlines some indisputable advances: primarily, the correlation between null subjects and the morphology of rich agreement in inflectional languages; and the correlation between referential null subjects and obligatory expletive null subjects in NSLs.

Contributions from empirical variationist studies

This phenomenon has been widely researched, and an extensive literature has been published within both the functionalist approach and the generative model, producing significant results in terms of the distribution of expressed subjects and null subjects in natural languages. Nonetheless, certain questions remain unresolved, as Ayres (2021) observes in the introduction to her doctoral thesis. One such question relates to the factors that favour the use of null subjects versus pronominal subjects. A large proportion of publications in the variationist field deal with varieties of Portuguese and Spanish, whether these are peninsular and other European varieties, or the geographical varieties located principally on the American continent. Other variationist studies have focused on heritage varieties spoken by emigrants living for example in Toronto (NAGY, 2015) and New York (OTHEGUY, ZENTELLA & LIVERT, 2007).

Within the variationist field, four recent studies provide a synthesis of the linguistic factors that have come to be considered those which condition the variable null subject (Table 5).

Table 5. Most widely studied linguistic factors in the variation of pronominal vs. null subject expression

Linguistic factors	Variables
Type of subject	<ul style="list-style-type: none"> • Null-argument subject, expletive, no referential content • Subject of a meteorological verb • Argument subject, but arbitrary or of an indefinite interpretation • Argument subject with definite reference
Type of clause	<ul style="list-style-type: none"> • Finite clause • Co-ordinate • Subordinate
Distance	<ul style="list-style-type: none"> • Adjacent • Non-adjacent
“priming” effect (effect of the antecedent)	<ul style="list-style-type: none"> • A null subject • a pronominal subject • a nominal subject
Verbal semantic content	<ul style="list-style-type: none"> • Mental activity or epistemics • Stative verbs • Dynamic verbs
Verbal morphology	<ul style="list-style-type: none"> • Simple past • Imperfect • Present
Person and number	<ul style="list-style-type: none"> • 1st-person singular • 2nd-person singular • 3rd-person singular • 1st-person plural • 2nd-person plural • 3rd-person plural
Animacy	<ul style="list-style-type: none"> • Animate • Inanimate
Coreferentiality	<ul style="list-style-type: none"> • With a subject • With an indirect object • With a direct object • With an oblique object

Sources: Bouchard (2018), Posio (2018), Torres Cacoullós & Travis (2019), Flores & Rinke (2020).

These factors include the **type of subject, grammatical category of person/number, coreferentiality, and the “priming effect”** or antecedent, all of which have been referred to above. However, there are also other factors related to the adjacency and processing of elements, to the verbal semantic and morphological content, and to the feature of animacy of the subject referent.

On extralinguistic and social factors

By contrast with the diverse range and relevance of structural and linguistic factors that have been tested in empirical (variationist) studies, factors of a social or classically extralinguistic nature (age, sex, level of education) seem to have a lesser incidence on the variations (Table 6).

Table 6. Most widely studied sociolinguistic factors for variations in pronominal subject / null subject expression

Varieties of Portuguese	<ul style="list-style-type: none"> • <i>Level of formal education:</i> Bouchard (2018:15), São Tomé Portuguese (use of null subjects: 0.55 primary school; 0.50 high school; 0.45 university) • <i>Gender and living away from the community:</i> Lucchesi (2009: 181), Afro-Brazilian Portuguese • none: Duarte & Reis (2018: 187, PB (36 CORPORAPORT)); Reis (2020) Madeira Island European Portuguese (Funchal, CORPORAPORT)
Varieties of Spanish	<ul style="list-style-type: none"> • none: Otheguy et al. (2007:778) Spanish-speaking residents of New York. • <i>age:</i> Orozco & Guy (2008): E. Barranquilla, Colombia; Alfaraz (2015): Santo Domingo, Dominican Republic; Lastra & Butragueño (2015): Mexico City (Mexico); Prada Pérez (2015), bilingual Catalan-Spanish speakers, Menorca, Spain. • <i>gender:</i> (Posio 2016)

Sources: Bouchard (2018), Lucchesi (2009), Duarte & Reis (2018), Reis (2020), Othegui et al. (2007), Orozco & Guy (2008), Alfaraz (2015), Lastra & Butragueño (2015), Prada Pérez (2015), Posio (2016).

In some studies, just one of the variables has a statistical effect – such is the case with the *age* variable in varieties of Spanish (OROZCO & GUY, 2008; LASTRA & BUTRAGUEÑO, 2015; ALFARAZ, 2015), or the *sex* variable in varieties of peninsular Spanish and peninsular European Portuguese (POSIO, 2016). In other studies, however, none of the three variables is shown to be statistically relevant, as will be discussed in more detail later in this article.

The production of null subjects in spoken varieties of Portuguese

In its traditional form of standard European Portuguese, Portuguese as a consistent Null Subject Language is marked by the possibility of the omission of the referential subject. In addition to a morphology of rich verbal agreement in which the verbal form already possesses affixes of Person/Number, standard European Portuguese exhibits a set of properties (LOBO, 2016: 562) that are illustrated by the examples given in (6) to (9):

(6) Possibility of referential null subjects

a. \emptyset Fomos à universidade.

b. \emptyset Conheço a Maria.

(7) Possibility of expletive null subjects

c. Choveu.

d. Parece que vai chover.

(8) Possibility of inversion of the order of subject-verb

e. Nasceu o meu sobrinho!

f. Floriram as cerejeiras em Trás-os Montes!

(9) Possibility of extraction of the subject at a distance

- g. Quem você/ocê disse que vai via
- h. Quem você/ocê acha que foi o vencedor

Semantic factors such as *animacy* must also be considered, because they take on great importance in varieties of Portuguese. In standard European Portuguese, there is a tendency for overt pronominal subjects to be [+animate] and for pronominal subjects that refer to [-animate] antecedents to be null (BARBOSA, DUARTE & KATO, 2005; LOBO, 2016):

- (10) a. O João leu que o *tsunami de 2004*_i foi o desastre natural mais mortal da história. No total, #*ele*_i tirou a vida a 230 000 pessoas.
- b. O João leu que *Jack o estripador*_i é um dos criminosos mais famosos de sempre. No total, *ele*_i tirou a vida a 5 pessoas.

Not all varieties of Portuguese observe the same pattern of properties. The literature on the differences between European Portuguese and Brazilian Portuguese is extensive. Spoken Brazilian Portuguese avoids leaving the position before the verb empty (KATO, 2002), as the following examples illustrate:

- (11) a. Ø Durmo cedo.
a'. *Eu durmo cedo.*
- b. Ø *Conserta sapato(s)*
b'. *Aqui conserta sapato.*
- c. Ø *Vira à esquerda.*
c'. *Cê vira à esquerda.*
- d. Ø *Dormem ali os meninos.*
d'. *Ali dormem os meninos*
- e. Ø *Chove em São Paulo.*
e'. *São Paulo chove.*
- f. Ø *Furou o pneu do Hamilton.*
f'. *O Hamilton furou o pneu.*
- g. Ø *Xinguei o cara.*
g. *Dá xinguei o cara.*

Table 7 shows some of the differences between European Portuguese and Brazilian Portuguese in relation to the possibility of the production of null subjects, considering the correlation of the subject with other properties such as the type of subject, coreferentiality, and the animacy or not of the subject, as well as other grammatical possibilities within these two varieties, as is the case in the production of more complex structures such as that of double subjects.

Table 7. Principal structural differences in the production of null subjects (vs. overt subjects) in EP and BP

	European Portuguese	Brazilian Portuguese
Omission of referential subjects	Ø Fomos à praia	A gente foi à praia / Nós fomos à praia.
Arbitrary null subjects	É assim que *(se) faz o doce Nesse hotel não *(se) pode entrar na piscina bêbado	É assim que faz o doce Nesse hotel não pode entrar na piscina bêbado.
Coreferentiality	O chefe _i disse ao amigo _j que <i>pro</i> _i precisava de descansar O chefe _i disse ao amigo _j que ele _j precisava de descansar	A Ana _i disse à Rosa _j que ela _{i/j} precisava de descansar O povo brasileiro _i acha que ele _i tem uma grave doença
Animacy	A casa _i virou um filme quando teve de ir abaixo.	A casa _i virou um filme quando ela _i teve de ir abaixo.
Double subjects	*A Clarinha ela cozinha que é uma maravilha.	A Clarinha ela cozinha que é uma maravilha.
Arbitrary pronominal subjects	Øexpl Deviam ensinar o amor às crianças. Øexpl Chove muito nessas florestas.	<i>Eles</i> deviam ensinar amor às crianças. <i>Essas florestas</i> chovem muito. (Kato e Duarte, 2014)

Source: Lobo (2016: 572).

By contrast with European Portuguese, Brazilian Portuguese reveals a tendency to complete the subject, whether it is an argument or referential, quasi-argument, or arbitrary. In this variety, pronominal referential subjects may also correspond to [-animate] referents, meaning that they are not subject to the semantic restrictions observed in European Portuguese. On the other hand, and contrary to European Portuguese, the subject in the main clause and the pronominal subject in the subordinate clause may be coreferential in Brazilian Portuguese. The duplication of the subject by a pronoun is possible in Brazilian Portuguese and not in standard European Portuguese. In both varieties, expletive subjects always correspond to null subjects (*Ele parece que *ele vai chover amanhã). With regard to arbitrary subjects, in Brazilian Portuguese there is the use of the generic null, as Kato & Duarte (2014) underline, and the disappearance of the indefinite *se* that is used in European Portuguese (GALVES, 1987). This

variant competes with other that have an expressed pronominal subject such as *a gente* or *você*, as in the examples (12) given by Kato & Duarte (2014: 9):

- (12) a. \emptyset não pode fumar aqui.
 b. *A gente* não pode fumar aqui.
 c. *Você* não pode fumar aqui.

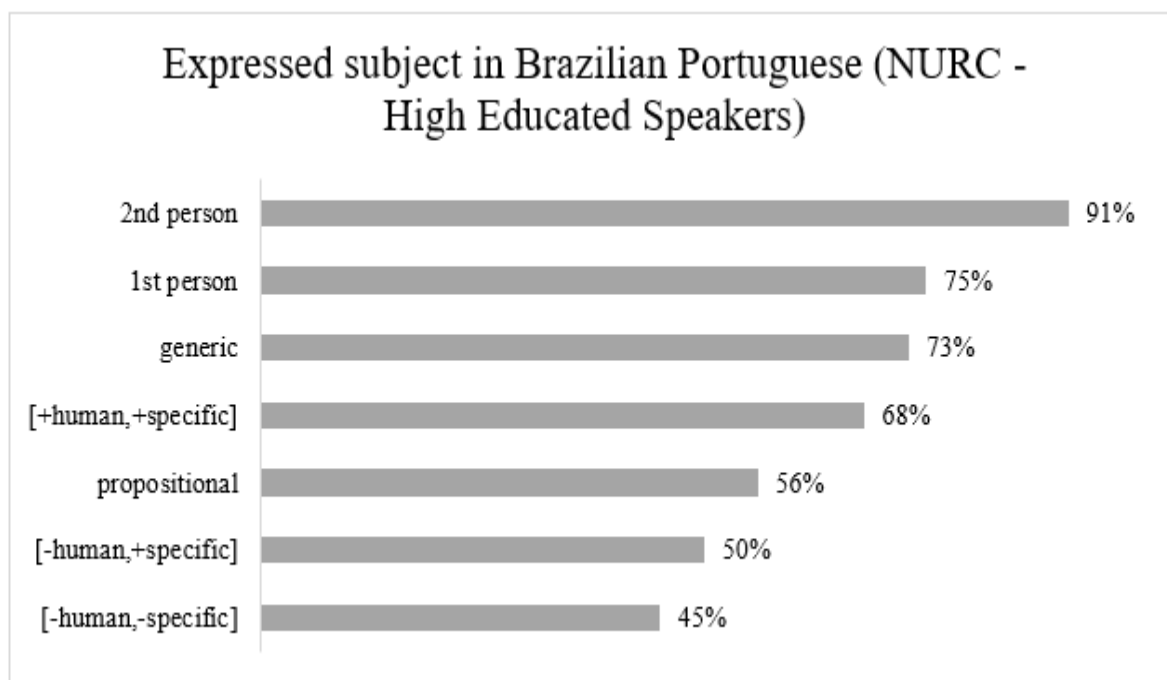
For Holmberg *et al* (2009), this is a construction found in *partial* Null Subject Languages, but not in consistent Null Subject Languages such as Italian and Spanish, which require the clitic *se*.

It is possible therefore to observe that Brazilian Portuguese has significantly reduced the contexts in which the null subject may be realised. Cyrino, Duarte & Kato (2000) observe on the basis of their work on syntactic specificities in the area of the subject and the direct object in Brazilian Portuguese that referentiality is very important for the realisation of the pronoun. To account for these specificities, they propose the concept of a referential hierarchy, in (i):

(i) **Referential hierarchy (Cyrino, Duarte & Kato, 2000: 39)**

null-argument	proposition	P3	P2 P1
		+human	±human
		±specific	
[-referential] < -----X-----X-----> [+referential]			

The pronominalisation of the subject is implemented when the conditions are in place for greater referentiality, in other words, in the 1st and 2nd persons, with [+human]; and it diminishes progressively with antecedents of lesser referentiality, first in the 3rd person [+human] followed by [-human]. Arbitrary [-referential] subjects offer greater resistance to pronominal omission and are therefore null, and in this they are consistent with the standard variety of European Portuguese. Graph 2, below, illustrates the way in which this referential hierarchy is applied in Brazilian Portuguese, using data from the Norma Urbana Oral Culta (NURC) of educated speech from Rio de Janeiro, first analysed by Duarte (1995) and referred to in Kato & Duarte (2014).

Graph 2. Expressed subjects (vs. null subjects) within the referential hierarchy of BP

Source: Kato & Duarte (2014: 5)

The differences between European Portuguese and Brazilian Portuguese in relation to the possibility of the production of null subjects are summarised here, and they lead to the conclusion that the grammar of the Brazilian variety of Portuguese has been the object of a process of change during which it has lost properties that are prototypical of a Null Subject Language such as European Portuguese. The hypothesis that BP can be considered a *partial* NSL, as proposed by Holmberg & Sheehan (2010) has not drawn a consensus. Kato (2020: 77) outlines four hypotheses for the direction of change underway in BP, which are that it will become: i) a non-Null Subject Language of the same type as French and English; ii) a *radical* Null Subject Language, like Chinese and Japanese; iii) a *partial* Null Subject Language such as Finnish; or iv) a “semi” Null Subject Language similar to Icelandic. Kato (2020: 89) reviews each of these hypotheses and ultimately argues for a framework in which a competition between grammars can be observed in Brazilian Portuguese – between the innovative, which is acquired as a core grammar, and the conservative, which is learned as a second grammar through formal instruction. According to the author, this core grammar would be similar to that of Icelandic, “com sujeitos referenciais que consistem em um paradigma de pronomes fracos livres como o inglês e o francês, mas com elementos nulos no lugar de sujeitos não referenciais” [“with referential subjects that consist in a paradigm of free weak pronouns like in English and French, but with null elements in the place of non-referential subjects”] (KATO, 2020: 89).

The corpora

The comparison of results is based on *corpora* comprising the findings of empirical research detailed in fourteen publications, listed in Table 8 below.

Table 8. Corpora of publications on the production of null subjects in spoken varieties of Portuguese

Varieties	Studies	Corpora
Afro-Brazilian Portuguese	Lucchesi, 2009	Corpus Vertentes
Dialectal Brazilian Portuguese	Myamoto & Kailer, 2019	Sample of the Atlas Linguístico do Brasil (ALiB)
Brazilian Portuguese (Rio de Janeiro)	Duarte, 1995; Soares, 2017;	NURC (Rio de Janeiro); Corpus Concordância ⁶ - Rio de Janeiro
European Portuguese (Lisbon)	Duarte & Reis, 2018; Duarte, 2020	Corpus Concordância (Lisbon)
Dialectal European Portuguese	Carvalho, 2009	Corpus Cordial-SIN
Madeira Island European Portuguese	Reis, 2020	<i>Corpus</i> Concordância (Funchal)
São Tomé Portuguese	Bouchard, 2018	<i>Corpus</i> Bouchard (L1, Bilingual)
Angolan Portuguese	Oliveira & Santos, 2007	<i>Corpus</i> Chavagne (2005)
Popular Angolan Portuguese	Teixeira, 2012	<i>Corpus</i> Teixeira (L1, Bilingual)
Mozambican Portuguese	Oliveira, 2016	<i>Corpus</i> Anna Jon-And (L2, Maputo)
Heritage Portuguese (Hamburg) (HP Hamburg)	Flores & Rinke, 2020	<i>Corpus</i> of Heritage Portuguese (1 st and 2 nd Generations - Hamburg)

Source: Elaborated by the author

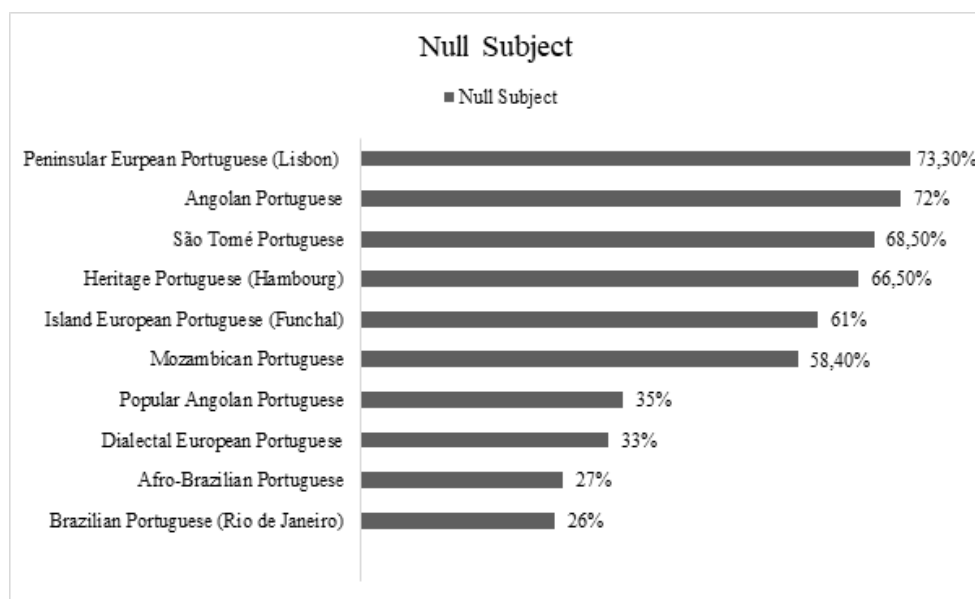
The selection seeks to bring together contributions that are focused on data related to spoken varieties of Portuguese. It covers geographical and social varieties, and different modes of acquisition of Portuguese (African, Brazilian, and European varieties, rural and urban varieties, L1 and also L2 with or without bilingualism, Heritage Languages). Research focused on data pertaining to written production was not included in this study, and discussions that centre almost exclusively on the theoretical discussion of the null subject parameter within the framework advocated by generative grammar were also excluded.

⁶ The Concordância *Corpus* currently forms part of the COPORAPORT project.

Results

The global results (Graph 3) show the configuration of a *continuum* on which, as expected, the peninsular variety of European Portuguese is located at one extreme that is characterised by a higher percentage of null subjects; at the opposite end, various varieties of Brazilian Portuguese are found, in all of which the percentage of null subjects is much smaller. Between these two broad varieties of European Portuguese and Brazilian Portuguese are located the Madeira Island variety of European Portuguese and various African varieties of Portuguese.

Graph 3. Variations in Null Subject expression in spoken varieties of Portuguese.

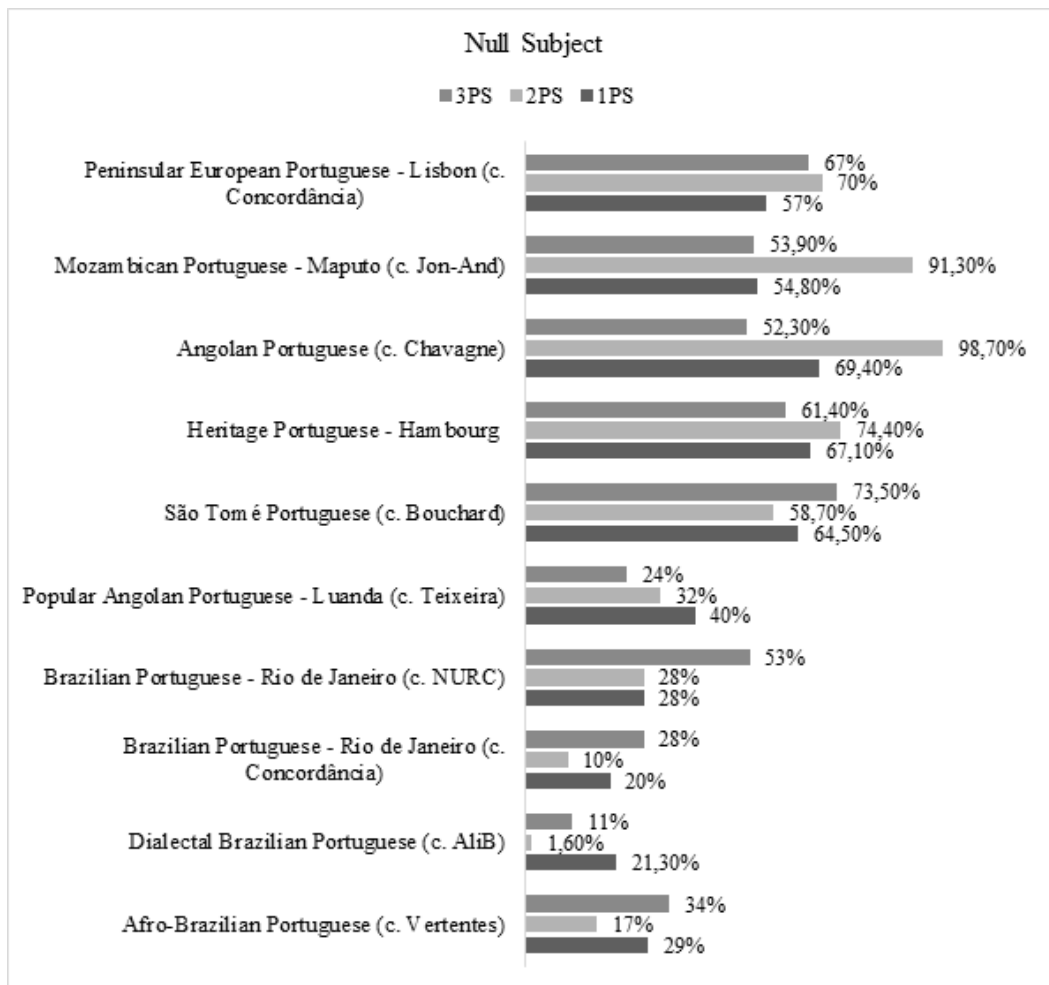


Sources: Peninsular European Portuguese (Duarte, 2000), Madeira Island European Portuguese (Reis, 2020); Dialectal European Portuguese, based on CORDIAL-SIN (Carvalho, 2009); Brazilian Portuguese (Duarte, 2000); Afro-Brazilian Portuguese (Luccchesi, 2009); São Tomé Portuguese (Bouchard, 2018), Angolan Portuguese (Oliveira & Santos, 2007), Popular Angolan Portuguese (Teixeira, 2012), Mozambican Portuguese (Oliveira, 2016); Heritage Portuguese (Germany) (Flores & Rinke, 2020).

Contrary to what might be expected, the results pertaining to the production of null subjects in the socially opposed varieties of Afro-Brazilian Portuguese (LUCCHESI, 2009) and educated Brazilian Portuguese (with a sample from the NURC, DUARTE, 1995; SOARES, 2017) give very similar outcomes of 27% and 26% respectively. Therefore, the change that has happened in BP must not be related to the level of formal education of the speakers, nor is their background related to language contact, as Lucchesi (2009: 176) observes. For Lucchesi, this situation of convergence arises from two developmental trajectories: that of the educated variety and the structural reorganisation that is provoked by the substitution of the pronouns *tu* and *nós* by *você* and *a gente*; and that of popular varieties, in which inflected morphemes of person and number are dropped in the irregular linguistic transmission that occurs through the contact of Portuguese with African languages.

The results of the uses of the null subject distributed across the 1PS, 2PS, and 3PS are presented in Graph 4 below:

Graph 4: Variations in Null Subject expression by PERSON/NUMBER (1PS, 2PS, and 3PS) in spoken varieties of Portuguese.



Sources: Afro-Brazilian Portuguese (*corpus* Vertentes, cf. Lucchese, 2004); Dialectal Brazilian Portuguese (*corpus* ALiB, Miyomoto & Kailer, 2019), Brazilian Portuguese – Educated Speakers (*corpus* NURC, Soares, 2017) Brazilian Portuguese (*corpus* Concordância, Duarte, 2020), Peninsular European Portuguese (*corpus* Concordância, Duarte, 2020), São Tomé Portuguese (Bouchard, 2018), Popular Angolan Portuguese (Teixeira, 2012), Mozambican Portuguese (*corpus* Jon-And, Oliveira, 2016) and Heritage Portuguese - Hamburg (Flores & Rinke, 2020).

There is a clear tendency towards the more frequent omission of the subject in the context of the 3PS in all varieties, although less so in the Popular Angolan variety (Luanda) and in the BP variety in the ALiB data. In these varieties, null subjects occur more frequently in the 1PS.

In relation to the other linguistic varieties, and as Bouchard (2018) observes, the studies selected do not present results that can be usefully compared because they make use of different variables and analytical methods. Some exceptions, however, are the recent studies by Duarte (2020) and Reis (2020) based on the Concordância *corpus*, which includes samples of Brazilian

Portuguese (Rio de Janeiro: Nova Iguaçu and Copacabana), a sample of Peninsular European Portuguese (Lisbon: Cacém and Oeiras), and a sample of Madeira Island European Portuguese (Funchal). The results presented in Chart 1 below account for the correlation between *animacy* and the production of null subjects.

Chart 1. Variations in Null Subject expression in the 3PS and its correlation with the semantic features of the Subject in spoken varieties of European Portuguese and Brazilian Portuguese.

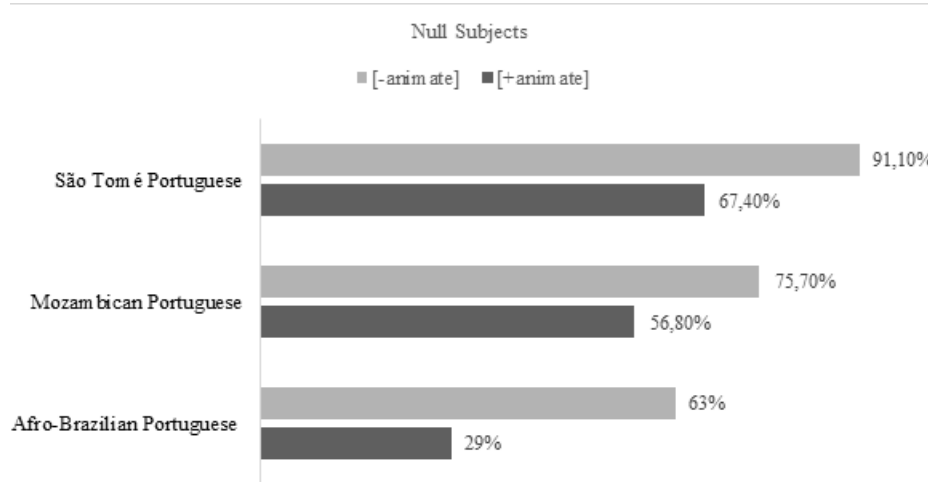
Features	Peninsular European Portuguese input 0.756			Madeira Island European Portuguese (Funchal) input 0.748			Brazilian Portuguese (Rio de Janeiro) input 0.248		
	N/T	%	P.R.	N/T	%	P.R.	N/T	%	P.R.
[-animate /-specific]	12/12	100%	-	12/12	100%	-	7/12	58%	0.863
[-animate /+specific]	137/142	96.5%	0.942	180/187	96%	0.919	73/173	42%	0.692
[+animate /-specific]	191/246	78%	0.562	155/194	80%	0.547	62/191	32.5%	0.555
[+animate /+specific]	308/559	55%	0.307	173/403	43%	0.229	189/803	23.5%	0.437
	<i>range</i>		<i>0.635</i>			<i>0.690</i>	<i>range</i>		<i>0.426</i>
	Log likelihood = -440.725 Significance = 0.000			Log likelihood = -350.925 Significance = 0.000			Log likelihood = -624.928 Significance = 0.000		

Sources: Peninsular European Portuguese and Brazilian Portuguese (Duarte, 2020) and Madeira Island European Portuguese (Reis, 2020)

It can be observed in these results that the referential subjects [-animate, +specific] and [-animate, -specific] in the two varieties of European Portuguese correspond to categorical (100% in both) and almost categorical (96.5% and 96%) null subjects, while this does not occur in Brazilian Portuguese. In this latter variety, a significant tendency towards the use of null subjects in these contexts (58% and 42%) can also be seen, although in a much smaller number. The most striking contrast is evident in contexts where the feature [+animate] is predominant: BP tends to use overt subjects (67.5% and 76.5%), while European Portuguese uses null subjects (78% and 55%). Contexts with the feature [-animate] offer the greatest resistance to the implementation of the forms of the subject expressed in Rio de Janeiro varieties of Brazilian Portuguese.

It is possible to compare in absolute terms the effect of animacy on the Afro-Brazilian variety and on two African varieties of Portuguese (Mozambican Portuguese and São Tomé Portuguese) in other studies.

Graph 5. Variations in Null Subject expression according to the variable animacy of the subject in three spoken varieties of Portuguese.



Sources: Afro-Brazilian Portuguese (Lucchesi 2009); Mozambican Portuguese (Oliveira, 2016, with the Jon-And corpus of Maputo, L2); São Tomé Portuguese (Bouchard, 2018).

The same pattern occurs across the three varieties: the feature [-animate] stimulates considerably the production of null subjects. However, only the São Tomé variety seems to come close to the European Portuguese data.

The contrasts between varieties of European Portuguese and Brazilian Portuguese may be less marked when we consider the data for Dialectal European Portuguese. Carvalho’s (2009) work has shown that the percentage of use of *null subjects* is not uniform in EP and that it produces figures that do not correspond to a classical or consistent Null Subject Language. In her study of dialectal (peninsular and island) varieties of EP, based on a sample from CORDIAL-SIN, the author presents percentages for pronominal subjects that are higher than those for null subjects. Only one locality does not conform to this pattern: Figueiró, in the district of Guarda. In this locality, the percentage for null subjects (64%) is superior to that for pronominal subjects (36%) (Chart 2).

Chart 2. [+specific] subjects in dialectal European Portuguese (data from CORDIAL-SIN)

Locality	Pronominal subject	Null subject	Total
Alcochete	48 (69%)	22 (31%)	70
Alvor	78 (76%)	25 (24%)	103
Cabeço de Vide	326 (92%)	29 (8%)	355
Camacha	85 (62%)	51 (38%)	136
Ponta da Garça	288 (93%)	22 (7%)	310
Serpa	19 (53%)	17 (47%)	36
Arcos	167 (65%)	90 (35%)	257
Outeiro	86 (55%)	70 (45%)	156
Figueiró	70 (36%)	123 (64%)	193
Covo	561(67%)	276 (33%)	837
	1728 (70%)	725 (30%)	2453

Source: Carvalho (2009: 134)

The most surprising result corresponds to the locality of Ponta da Garça on the island of São Miguel (Azores), with 93% of expressed subjects – a higher percentage than that for the BP variety that has the highest percentage for completion of the subject in the publications selected for this study (73%).

In relation to the effect of the grammatical person/number variable on the use of null subjects, the dialectal data for EP reveal a certain discrepancy with regard to those expressed in the varieties of EP that are included in the *Concordância Corpus*, as attested in Chart 3.

Chart 3. PERSON/NUMBER variable and subject expression in dialectal European Portuguese data from CORDIAL-SIN

	1PS (eu)	2PS (tu)	3PS (ela)	3S/1P (a gente)	1PP (nós)	3PP (eles)	Total
Null subjects	417 27%	20 22%	209 42%	10 11%	20 37%	46 57%	721 30%
Overt subjects	1132 73%	71 78%	296 58%	83 89%	34 63%	35 43%	1651 70%
	1548	91	505	93	54	81	2372

Source: Carvalho (2009: 145)

Subject expression produces very high values in the 1PS (73%) and in the 2PS (78%). This tendency is counter to that shown in Duarte's (2020) results, which are more in accordance with the properties of a Null Subject Language, with percentages for subject expression of 43% and 30% respectively.

Another result worthy of mention here comes from an island variety of EP, that also contrasts with the properties attributed to urban varieties of European Portuguese and Brazilian Portuguese. A study by Casalta (2017), based on data taken from the locality of Camacha on the island of Porto Santo (Madeira Archipelago) and made available in the CORDIAL-SIN *corpus* shows that in 14 arbitrary subjects there are 10 pronominal subjects. The examples at (13) are interpreted as arbitrary pronominal subjects [+animate] because they represent the indefinite referent *people*.

- (13) a. *eles agora costumam de lavar a vinha.* (PST02)
 b. *Aqui, eles andaram fazendo e deram com água.* (PST07)

These examples diverge from the standard variety of European Portuguese. In that variety, arbitrary subjects in the 3PP are categorically null and an expressed pronoun in the 3PP can only have a specific referential interpretation.

Finally, the construction of the double subject should also be mentioned. This is a complex structure that combines the pronoun *a gente* with the clitic *se*. It is very common in Madeira Island European Portuguese, whether in the urban variety (Funchal) or in the rural varieties (HENRIQUES, 2019), and it is also found in peninsular varieties of European Portuguese. The

double subject expressed by *a gente se* has an arbitrary interpretation and a complex nature, because its referential reach can oscillate between the generic [-specific, +impersonal] and the deictic [+specific, -impersonal], as in the following examples:

- (14) *generic*: Havia um arraial ali que **a gente** chama-se “Maio”.
definite: **Bordava-se** o sábado, só o domingo é que ia-se p’o caminho brincar.
specific: **A gente** era-se garotos pequenos.
deictic: Ah, pega-me às costas. **A gente** chega-se lá num instante.

(HENRIQUES, 2019: 57-58)

This construction is found in non-standard varieties of European Portuguese and is more frequent in the variety spoken on island of Madeira. It constitutes a variant of the double subject construction. It is rooted in European Portuguese because it maintains the clitic *se*, contrary to what is observed in the varieties of Brazilian Portuguese:

- (15) a. No Brasil, não se usa mais saia (European Portuguese)
b. No Brasil, não \emptyset usa mais saia. (Brazilian Portuguese)

As Barbosa, Duarte & Kato (2005) recognise, in Brazilian Portuguese other variants of the overt subject are used, expressed by *a gente* and *você*, as shown in (16):

- (16) a. E se *a gente* falar que não tem? (KATO & TARALLO, 1986: 347)
b. *Você*, no Canadá, *você* pode ser o que *você* quiser. (BARBOSA, DUARTE & KATO, 2005: 19)

Furthermore, in Mozambican Portuguese spoken in Maputo, it is possible to observe indeterminate subjects that are similar to those in the Brazilian Portuguese data: using the third-person verb without *se* and without any special marker (17), and with the use of the pronoun *você* (18):

- (17) a. Às vezes há dias que não \emptyset consegue nada. (MZ17)
b. Aqui agora já não \emptyset faz mais isso. (MZ4)
- (18) a. Só quando *você* tem aquela classe pequena, mas nessa classe pequena *você* aproveitava. (MZ12)
b. Quando *você* vai para a escola, *você* não vai aprender nada. (MZ12)

(Oliveira, 2016: 67)

Social variables have not merited much attention in existing research. In many works, information pertaining to these variables is not made available. The *education level* variable was found to be statistically relevant in the work of Bouchard (2017) on the African variety of STP; the *age* variable in the MP variety (OLIVEIRA, 2016), and *sex* in the sample of Afro-Brazilian Portuguese analysed by Lucchesi (2009).

Table 9. Social variables and the production of null subjects in spoken varieties of Portuguese.

	age	sex	level of education	lived outside the community	contact with other languages
Afro-Brazilian Portuguese	No	Male 0.52	X	Never 0.53	X
Dialectal Brazilian Portuguese (<i>Corpus</i> ALiB)	No	No	No	X	X
Brazilian Portuguese (<i>Corpus</i> Concordância Rio Janeiro)	No	No	No	X	X
European Portuguese (<i>Corpus</i> Concordância Lisbon)	No	No	No	X	X
Dialectal European Portuguese (CORDIAL-SIN) *	X	X	X	X	X
São Tomé Portuguese (<i>Corpus</i> Bouchard)	-	-	Primary 0.55 Secondary 0.50	-	**
Angolan Portuguese (<i>Corpus</i> Chavaigne)*	X	X	X	X	X
Popular Angolan Portuguese (<i>Corpus</i> Teixeira)	X	X	X	X	X
Mozambican Portuguese (<i>Corpus</i> Jon-And)	35-5 9 0.55				
Heritage Portuguese (<i>Corpus</i> Hamburg, Germany) ***	X	X	X	X	X

* Profile of informants: illiterate, elderly, born and always lived in the same locality

** Level of education variable

*** Forms a single extralinguistic variable: 1st generation vs 2nd-generation speakers

X information unavailable/not published.

Sources: Afro-Brazilian Portuguese (Lucchesi, 2009); Mozambican Portuguese (Oliveira, 2016, with the Jon-And corpus (Maputo, L2); São Tomé Portuguese (Bouchard, 2018); Dialectal European Portuguese (Carvalho, 2009, with data from CORDIAL-SIN); Brazilian Portuguese (Duarte, 2020, *Corpus* “concordância-RJ”); Dialectal Brazilian Portugueses (Miyomoto & Kailer, 2019, with data from) ALiB (Atlas Linguístico do Brasil); Angolan Portuguese (Oliveira & Santos, 2007 – corpus Chavaigne); Popular Angolan Portuguese (Teixeira, 2012); Heritage Portuguese – Hamburg (Flores & Rinke, 2020).

Final Considerations

Following this examination of the production of null subjects in spoken varieties of Portuguese, the difficulties in making a comparison of the data pertaining to these varieties are patently clear. The research methodology used in each of the different studies selected and analysed here is variable, and this makes the task of comparing the results difficult and at times impossible.

It has only been possible to compare the different varieties according to the total percentages of null subjects, and their distribution across the grammatical category of Person (1st, 2nd, 3rd). In these two comparisons, a gradual increase in the percentages of null subjects is observed in the direction of the varieties of European Portuguese spoken in Lisbon (Oeiras and Cacém), when traced from the starting point of the Afro-Brazilian variety and other varieties of Brazilian Portuguese at the opposite end of the scale. The hypothesis that arises from this exercise is that this sphere is barely permeable to the effects resulting from contact with other languages. Whether in the situation of the variety of Heritage Portuguese spoken in Hamburg, Germany, where the dominant language of German – an expletive Null Subject Language – is in contact with the consistent Null Subject Language of Portuguese, or in the situation of African varieties which have to a greater or lesser extent become native varieties of Portuguese, the percentages of null subjects are consistent with the typological determination of European Portuguese.

The analysis of other structural factors such as the correlation of the production of null subjects and the semantic feature of *animacy* has only been partially possible. This illustrates the principal difficulty faced in this study: the differences in methodological approaches outlined above and observed by Bouchard. As emphasised in the literature, the [-animate] feature is realised categorically by null subjects in varieties of European Portuguese and almost categorically in African varieties such as São Tomé Portuguese. In Brazilian Portuguese, this semantic feature presents the strongest resistance to change in the direction of overt subject expression.

In general, the low significance of social / extralinguistic factors should be pointed out. Indirectly, and although it does not seem to have an effect on the varieties of Brazilian Portuguese, the *education level* factor proves to have a greater impact on varieties of European Portuguese. In European Portuguese, dialectal varieties that are typically rural and produced by informants with a low level of formal education reveal general rates of null subjects far inferior to those of the social urban varieties of the Lisbon region, which may indicate the necessity for a more detailed examination of European Portuguese data as a complex object in which various structural conditions interact. Furthermore, the convergence in uses of null subjects across Afro-Brazilian varieties and more highly educated varieties of Brazilian Portuguese (Rio de Janeiro) seems pertinent and worthy of deeper reflection.

Finally, it is worth highlighting the role of dialectal diversity observed in the linguistic area of Portuguese, and especially in the sphere of European Portuguese, which occurs in similar fashion to what happens in other romance languages such as Italian. The dialectal data for European Portuguese provide empirical counter-evidence against the parameters and typologies that are based almost exclusively on the possibility or otherwise of the production of null subjects. In this field, it would perhaps be preferable to face the possibility of recourse to the concept of *gradation* in variations, and to include it conceptually within the framework of parametric change, as Oliviéri (2004: 119-120) and Hinzelin & Kaiser (2012: 259-260) advocate in respect of the structural variations observed in the romance family.

References

ALFARAZ, G. G. *Variation of Overt and Null Subject Pronouns in the Spanish of Santo Domingo*. Washington, D.C.: Georgetown University Press, p. 3-16, 2015.

ALiB. <https://alib.ufba.br>.

ARENDS, J. Simple grammars, complex languages, *Linguistic Typology*, 5, p.180-182, 2001.

ARIEL, M. *Accessing Noun-Phrase Antecedents*. London: Routledge, 1990.

AUGER, J. Les pronoms clitiques sujets en picard: une analyse au confluent de la phonologie, de la morphologie et de la syntaxe. *Journal of French Language Studies*, 13, p.1–22., 2003. <https://doi.org/10.1017/S0959269503001066>.

AYRES, M. R. *Contextos licenciadores de sujeitos nulos em português brasileiro*. PhD Thesis. Instituto de Letras da Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 2021.

BARBOSA P.; DUARTE, M.E.L.; KATO, M. Null subjects in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4, p. 11–52, 2005.

BENINCÀ, P. Subject clitics and particles in Provençal, *Probus*, 26(2), p. 183-215, 2014.

BOUCHARD, M-E. Subject Pronoun Expression in Santomean Portuguese. *Journal of Portuguese Linguistics*, 17 (5), pp. 1–29, 2018. DOI: <https://doi.org/10.5334/jpl.19>

CAMACHO, J. A. *Null Subjects*. Cambridge: Cambridge University Press, 2013.

CAMERON, R. *Pronominal and Null Subject Variation in Spanish: Constraints, Dialects, and Functional Compensation*. Unpublished Ph.D. thesis, University of Pennsylvania, Philadelphia, PA, USA, 1992.

CAMERON, R.; FLORES-FERRÁN, N. Perseverance of subject expression across regional dialects. *Spanish in Context*, 1, p.41–65, 2004. DOI: <https://doi.org/10.1075/sic.1.1.05cam>.

CARDINALETTI, A. Subjects and clause structure. In: HAEGEMAN, L. (ed.) *The new comparative syntax*. London: Longman, 1997, p.33–63.

CARDINALETTI, A.; STARKE, M. The typology of structural deficiency: A case study of the three classes of pronouns. In: RIEMSDIJK, H. van (ed.), *Clitics in the languages of Europe*. New York: Mouton de Gruyter, 1999, p. 145-233.

CARMINATI, M.N. *The processing of Italian subject pronouns*, 2002. *Doctoral Dissertations Available from Proquest*. AAI3039345. <https://scholarworks.umass.edu/dissertations/AAI3039345>

CARVALHO, G.A. *A realização do sujeito na fala do português europeu*. PhD Thesis. Universidade Estadual Paulista, Faculdade de Ciências e Letras de Araraquara, Brazil, 2009.

CASALTA, M. *Análise de dados do CORDIAL-SIN O Sujeito de terceira pessoa do plural em PST e UNS*. Trabalho final de Dialetologia. Lisbon: FLUL, 2017. *Análise de dados do CORDIAL-SIN O Sujeito de terceira pessoa do plural em PST e UNS* (researchgate.net).

MARTINS, A. M. (coord.) [2000-]. CORDIAL-SIN: Corpus Dialectal para o Estudo da Sintaxe / Syntax-oriented Corpus of Portuguese Dialects. Lisboa, Centro de Linguística da Universidade de Lisboa. URL: <http://www.clul.ulisboa.pt/en/10-research/314-cordial-s>

CHOCIEJ, J. *Polish null subjects: English influence on heritage Polish in Toronto*, Unpublished ms, Department of Linguistics, University of Toronto, 2011.

CHOMSKY, N. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

CYRINO, S.; DUARTE, M.E.L.; KATO, M. A. Visible subjects and invisible clitics in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E.V. (ed.), *Brazilian Portuguese and the Null Subject Parameter*. Vervuert: Frankfurt am Main, 2000, p.55-74.

D’ALESSANDRO, R. *The Null Subject Parameter Where are we and where are we headed*, 2014.

D’ALESSANDRO, R., Null Subjects. In: FÁBREGAS, A., MATEU, J. e PUTNAM, M. (ed.). *Contemporary Linguistic Parameters*. London: Bloomsbury Academic, 2015, p. 201-226.

DE CROUSAZ, I.; SHLONSKY, U. The Distribution of a Subject Clitic Pronoun in a Franco-Provençal Dialect and the Licensing of Pro, *Linguistic Inquiry*, XXXIV/3, 2003, p. 413-442.

DIÉMOZ, F. *Morphologie et syntaxe des pronoms personnels sujets dans les parlers francoprovençaux de la Vallée d’Aoste*. Tübingen: A. Francke, 2007.

DRYER, M. S. Expression of pronominal subjects. IN: DRYER, M.S.; HASPELMATH, M. (ed.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013.

DRYER, M.S.; HASPELMATH, M. (ed.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013 (Accessed on 18 Dec 2015).

DUARTE, M. E. L., A remarcação em curso no valor do parâmetro do sujeito nulo. *Cuadernos de la ALFAL*, 12 (2), p. 71-99, 2020.

DUARTE, M. E. L., Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil In: ROBERTS, I., e KATO, M. A. (ed.) *Português brasileiro: Uma viagem diacrônica. Homenagem a Fernando Tarallo*. Editora da Unicamp, Campinas, Brazil, 1993, p. 107–129.

DUARTE, M. E.L. , Sociolinguistics as a powerful tool to follow the course of a parametric change, *University of Pennsylvania Working Papers in Linguistics*, Vol. 25 (2) , Article 3. 2020. Available at: <https://repository.upenn.edu/pwpl/vol25/iss2/3>

DUARTE, M.E.L. The loss of the avoid pronoun principle in Brazilian Portuguese. In: KATO, M.A.; NEGRÃO, E.V. (ed.), *Brazilian Portuguese and the null subject parameter*. Frankfurt am Main: Vervuert Verlag, 2000, p. 17–36.

DUARTE, M.E.L.; REIS, E. P. R. Revisitando o sujeito pronominal vinte anos depois. *ReVEL* 16(30), p.173–197, 2018.

DUARTE, M. E. L. *A Perda do Princípio ‘Evite Pronome’ no Português Brasileiro*. PhD thesis. Universidade de Campinas: São Paulo, Brazil, 1995.

FIGUEIREDO SILVA, M.C., O estado da arte dos estudos sobre sujeitos nulos, posições de sujeito e marcas flexionais. *Revista Linguística*, 13(2), p. 191-211, 2017. DOI: <http://dx.doi.org/10.31513/linguistica.2017.v13n2a13510>

FLORES, C. ; RINKE, E., Factors Constraining Subject Expression in European Portuguese Spoken in Hamburg. A Bi-Generational Corpus Investigation, *Discours*, 26 | 2020, <http://journals.openedition.org/discours/10648>.

GALVES, C.A sintaxe do português brasileiro. - *Ensaio de Linguística* 7, 1987.

GILLIGAN, G.M. *A cross linguistic approach to the pro-drop parameter*. PhD dissertation: University of Southern California, 1987.

HASPELMATH, M. Parametric versus functional explanations of syntactic universals, In: BIBERAUER, T. (ed). *The Limits of Syntactic Variation*, Amsterdam: Benjamins, 2008, p. 75-107.

HEAP, D. *La variation grammaticale en géolinguistique: Les pronoms sujet en roman central*. München: Lincomm Europa, 2000.

HELASVUO, M-L. Agreement or crystallization: Patterns of 1st and 2nd person subjects and verbs of cognition in Finnish conversational interaction. *Journal of Pragmatics* 63, p. 63–78, 2014.

HENRIQUES, Y. *Particularidades morfossintáticas em variedades rurais do português falado na Ilha da Madeira*. Master thesis, University of Zurich, Switzerland, 2019.

HINZELIN, M-O. ; KAISER, G. A. Le paramètre du sujet nul dans les variétés dialectales de l'occitan et du francoprovençal, In: BARRA-JOVER, M (dir.), *Études de Linguistique gallo-romane*. Vincennes: Université de Vincennes, p. 247-260, 2012.

HOLMBERG, A. Null subject parameters. In: BIBERAUER, T., HOLMBERG, A., ROBERTS, I. e SHEEHAN, M., (ed.). *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press, 2010, p. 88–124.

HOLMBERG, A.; SHEEHAN, M. Control into finite clauses in partial null-subject languages, In: BIBERAUER, T., HOLMBERG, A., ROBERTS, I.; SHEEHAN, M., (ed.). *Parametric variation: null subjects in minimalist theory*. Cambridge: Cambridge University Press, 2010, p.125-152.

HOLMBERG, A; NAYUDU, A; SHEEHAN, M. Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, 63(1), p. 59-97, 2009.

JAEGGLI, O.; SAFIR, K. (org.). *The Null Subject Parameter*. Kluwer Academic Publishers, 1989.

JIA, L.; BAYLEY, R. Null pronoun variation in Mandarin Chinese. *University of Pennsylvania Working Papers in Linguistics* 8(3), p.103–116, 2002

KAISER, G. A.; HACK, F. M. Sujets et sujets nuls en romanche, In: ILIESCU, M. *et al.* (ed.). *XXVe CILPR Congrès International de Linguistique et de Philologie Romanes, Tome 7, Section 15*, Berlin: De Gruyter Mouton, 2010, p. 83-91.

KATO, M. A.; DUARTE, M. E. L. Restrições na distribuição de sujeitos nulos no Português Brasileiro. *Veredas* 18(1), p. 1–24, 2014.

KATO, M.A. A evolução da noção de parâmetros. *D.E.L.T.A.*, 18:2, p. 309-337, 2002.

KATO, M.A. As controvérsias sobre o sujeito nulo no português brasileiro, In: CASTAGNA, V.; QUAREZEMIN, S. (org.). *Travessias em língua portuguesa Pesquisa linguística, ensino e tradução*. SAIL 16, Edizioni Ca'Foscari, 2020, p. 74-92. DOI 10.30687/978-88-6969-461-5/004.

KATO, M.A.; DUARTE, M.E.L. Restrições na distribuição de sujeitos nulos no Português Brasileiro, *Veredas*, 18(1), 2014. 24970-Texto do artigo-97931-1-10-20160616 (1).pdf

LABOV, W. Quantitative reasoning in linguistics. In: AMMON, U.; DITTMAR, N.; MATTHEIER, K. J.; TRUDGILL, P. (ed.). *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society/, Ein internationales Handbuch zur Wissenschaft von Sprache und Gesellschaft*, vol. 1, Berlin & New York: Mouton de Gruyter, 2005, p. 6–22.

LASTRA, Y.; BUTRAGUEÑO, P. M. Subject pronoun expression in oral Mexican Spanish. In: CARVALHO, A.M.; OROZCO, R.; SHIN, N. L. (ed.), *Subject pronoun expression in Spanish: A cross-dialectal perspective*, Washington, DC: Georgetown University Press, 2015, p. 39–57.

LEE, D-Y.; YONEZAWA, Y. The role of the overt expression of first and second person subject in Japanese. *Journal of Pragmatics* 40(4), p.733–767, 2008

LOBO, M. Sujeito nulo: Sintaxe e interpretação. In: RAPOSO, E.B.P.; NASCIMENTO, M.F.B. do; MOTA, M.A.C.da; SEGURA, L.; Mendes, A. (ed.) *Gramática do Português*. Lisbon: Fundação Calouste Gulbenkian, 2013, p. 2309-2335.

LOBO, M. Sujeitos nulos: gramática do adulto, aquisição de l1 e variação dialetal. In: MARTINS, A.M.; CARRILHO, E. (eds). *Manual de linguística portuguesa*. Berlin: DE GRUYTER, 2016, p. 559-561. <https://doi.org/10.1515/9783110368840-023>

LOBO, M.; MARTINS, A. M. Subjects. In: DUFTER, A.; STARK, E. (eds). *Manual of Romance Morphosyntax and Syntax*. Berlin: De Gruyter, 2017, p. 27-88.

LUCCHESI, D. A realização do sujeito pronominal. In: LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. (org.). *O português afro-brasileiro*. Salvador: EDUFBA, 2009, p. 165-183.

MADEIRA, A.; FIÉIS, A.; TEIXEIRA, J. Microvariação na resolução de sujeitos pronominais: português europeu vs. Italiano. *Revista da Associação Portuguesa de Linguística*, Nº8, p. 155-167, 2021. DOI: <https://doi.org/10.26334/2183-9077/rapln8ano2021a11>

MIESTAMO, M.; SINNEMÄKI, K.; KARLSSON, F. *Language Complexity: Typology, Contact, Change*. Amsterdam: John Benjamins Publishing Company, 2008.

MONTALBETTI, M. *After Binding Properties. On the Interpretation of Pronouns*. PhD thesis, Massachusetts, MA, USA: MIT, 1984.

MYAMOTO, A. M.; KAILER, D. Sujeito nulo no português brasileiro: variação estável ou mudança em curso? *A Cor das Letras* 20(1), p. 24-32, 2019. DOI: 10.13102/cl.v20i1.4622

NAGY, N. A sociolinguistic view of null subjects and VOT in Toronto heritage languages, *Lingua*, 164, p. 309-327, 2015.

NAGY, N.; AGHDASI, N.; DENIS, D.; MOTUT, A. Null subjects in heritage languages: Contact effects in a cross-linguistic context. *University of Pennsylvania Working Papers in Linguistics*, Vol.17 (2), Selected Papers from NWAV 39, 2011.

NURC. <https://nurcrj.lettras.ufrj.br>.

OH, S-Y, Overt reference to speaker and recipient in Korean. *Discourse Studies* 9(4), p. 462–492, 2007. <https://doi.org/10.1177%2F1461445607079163>

OLIVEIRA, K. R. S. *Sujeito nulo ou sujeito preenchido? Uma análise da variedade falada pela comunidade quilombola Rio das rãs – Bahia*, dissertação de mestrado, Vitória da Conquista., Bahia, Brazil, 2020.

OLIVEIRA, M.; SANTOS, E. Pronomes nulos na posição de sujeito no português de Angola – um estudo preliminar. *Filol. lingüíst. port.*, 9, p. 85-101, 2007.

OLIVEIRA, V.M.S. *A expressão do sujeito no português de Moçambique*, Masters dissertation, UNICAMP, Brazil, 2016.

OLIVIÉRI, M. Paramètre du sujet nul et inversion du sujet dans les dialectes italiens et occitans. *Cahiers de Grammaire* 29, p. 105-120, 2004.

OLIVIÉRI, M.; LAI, J-P.; HEAP, D. Partial subject paradigms and feature geometry in Northern Occitan dialects. In: PERPIÑÁN, S.; HEAP, D.; MORENO-VILLAMAR, I.; SOTO-COROMINAS, A. (ed.). *Romance Languages and Linguistic Theory 11: Selected papers from the 44th Linguistic Symposium on Romance Languages (LSRL)*, London, Ontario, 2017, p. 157-167. <https://doi.org/10.1075/rllt.11.07oli>

OROZCO, R. Pronominal variation in Colombian Costeño Spanish. In: CARVALHO, A.M.; OROZCO, R.; SHIN, N.L. (ed.). *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Washington, DC: Georgetown University Press, 2015, p. 17–37.

OROZCO, R.; GUY, G. El uso variable de los pronombres sujetos: ¿Qué pasa en la costa Caribe colombiana?, In: WESTMORELAND, M.; THOMAS, J. *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics*, Somerville, MA: Cascadilla, 2008, p. 70-80.

OROZCO, R.; HURTADO, L.M. A variationist study of subject pronoun expression in Medellín, Colombia. *Languages* 6: 5. <https://doi.org/10.3390/languages601005>, 2021.

OTHEGUY, R.; ZENTELLA, A.C. *Spanish in New York: Language contact, dialect levelling, and structural continuity*. Oxford: Oxford University Press, 2012.

OTHEGUY, R.; ZENTELLA, A.C.; LIVERT, D. Language and dialect contact in Spanish of New York: Toward the formation of a speech community. *Language* 83(4), p. 770–802, 2007.

POLETTI, C., *La sintassi del soggetto nei dialetti italiani settentrionali*. Padova: Unipress, 1993.

POSIO, P. Properties of Pronominal Subjects. In: GEESLIN, K. L. (ed.), *The Cambridge Handbook of Spanish Linguistics*. Cambridge: Cambridge University Press, 2018, p. 286-306.

POSIO, P. Subject expression in grammaticalizing constructions: The case of *creo* and *acho* ‘I think’ in Spanish and Portuguese. *Journal of Pragmatics* 63, p. 5–18, 2014.

POSIO, P. You and we: Impersonal second person singular and other referential devices in Spanish sociolinguistic interviews. *Journal of Pragmatics*, 99, p. 1–16, 2016.

POSIO, P. Subject pronoun usage in formulaic sequences: Evidence from Peninsular Spanish In: CARVALHO, A.M.; OROZCO, R.; SHIN, N.L. (ed.). *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Washington, DC: Georgetown University Press, 2015, p.59–78.

POSIO, P. The expression of first-person-singular subjects in spoken Peninsular Spanish and European Portuguese: Semantic roles and formulaic sequences. *Folia Linguistica* 47 (1), p. 253-292, 2013.

PRADA PÉREZ, A. de First person singular subject pronoun expression in Spanish in contact with Catalan. In: CARVALHO, A.M.; OROZCO, R.; SHIN, N.L. (ed.). *Subject pronoun expression in Spanish: A cross-dialectal perspective*. Washington, DC: Georgetown University Press, 2015, p. 121-142.

VERTENTES. <http://www.vertentes.ufba.br/home>.

REIS, E.P. R. *Atravessando o oceano: a expressão do sujeito pronominal de referência definida no português europeu insular*. Monography. Rio de Janeiro: UFRJ, Brasil, 2020.

RIZZI, L. On the status of subject clitics in Romance. In: JAEGGLI, O.; SILVA-CORVALÁN, C. (eds.). *Studies in Romance linguistics*. Dordrecht: Foris, 1986, p. 391-419.

RIZZI, L. *Issues in Italian syntax*. Dordrecht: Foris, 1982.

SAMPSON, G.; GIL, D.; TRUDGILL, P. (ed.). *Language Complexity as an Evolving Variable*. Oxford; New York: Oxford University Press, 2009.

SAVOIA, L. M.; MANZINI, R. Les clitiques sujets dans les variétés occitanes et francoprovençales italiennes. *Corpus*, 9, p. 165-189, 2010.

SAVOIA, L.M.; MANZINI, M. R. Lexicalization of 3rd person object clitics: clitic enclisis and clitic drop. In: D’ALESSANDRO, R.; LEDGEWAY, A.; ROBERTS, I. (ed.). *Syntactic Variation. The Dialects of Italy*. Cambridge: Cambridge University Press, 2010, p. 102–118.

SHEEHAN, M. Subjects, null subjects, and expletives. In: FISCHER, S.; GABRIEL, C. (ed.). *Manual of Grammatical Interfaces in Romance*. Berlin: De Gruyter, 2016, p. 329-362. <https://doi.org/10.1515/9783110311860-014>

SHIN, N. L. Grammatical complexification in Spanish in New York: 3sg pronoun expression and verbal ambiguity. *Language Variation and Change* 26(3), p. 303–330, 2014.

SIBILLE, J. *La passion de saint André, édition critique suivie d’une étude linguistique comparée*. Paris: Honoré Champion, 2007.

SIBILLE, J. Les clitiques sujets dans le parler occitan de Chiomonte et des Ramats (Italie). *Revue de Linguistique Romane* 76, 401–435, 2012.

SILVA-CORVALÁN, C. Otra mirada a la expresión del sujeto como variable sintáctica. In; MORENO, F. F., et. Al. (ed.), *Lengua, variación y contexto: Estudios dedicados a Humberto López Morales*, vol. 2. Madrid: Arco Libros, 2003, p. 849–860.

SOARES, E. C. *Anaphors in discourse: anaphoric subjects in Brazilian Portuguese*. PhD Thesis. University Sorbonne Paris City, Paris, France, 2017.

SOARES, E.C.; MILLER, P.; HEMFORTH, B. O Efeito de Características Semânticas e Discursivas no Uso de Sujeitos Nulos e Pronominais - Um Estudo Quantitativo dos Sujeitos de Terceira Pessoa no Português do Brasil. *D.E.L.T.A.*, 36 (1), p. 1-38, 2020. <https://doi.org/10.1590/1678-460X2020360107>

TEIXEIRA, E. P., A representação do sujeito pronominal no português popular angolano, *PAPIA*, 22 (1), p. 141-159, 2012.

TORIBIO, J. A., Setting parametric limits on dialectal variation in Spanish. *Lingua* 110, p. 315–341, 2000.

TORRES CACOULOS, R.; TRAVIS, C. E. Variationist typology: Shared probabilistic constraints across (non-)null subject languages. *Linguistics* 57, p. 653–92, 2019.

TRAVIS, C.E. Genre effects on subject expression in Spanish: priming in narrative and conversation. *Language Variation and Change* 19 (2), p.101-135, 2007.

TRAVIS, C.E.; TORRES CACOULOS, R. What do subject pronouns do in discourse? Cognitive, mechanical and constructional factors in variation. *Cognitive Linguistics* 23(4), p. 711–748, 2012.

VIEIRA, S. R.; BRANDÃO, S. F. CORPORAPORT: Variedades do Português em análise. Rio de Janeiro: Faculdade de Letras-UFRJ. www.corporaport.lettras.ufrj.br.



PRONOMINAL SUBJECT IN ALAGOAS: A CASE OF CHANGE IN PROGRESS

O SUJEITO PRONOMINAL EM ALAGOAS: UM CASO DE MUDANÇA EM PROGRESSO

Maria Auxiliadora da Silva Cavalcante¹

ABSTRACT

Several studies demonstrate that Brazilian Portuguese is in a process of change in relation to the way in which pronominal subjects are carried out. There is a clear preference for full achievements over null subjects. However, these studies are mostly based on speech samples from the Southeast Region, which does not invalidate the research, but leaves room for some questions, especially if we take into account the large Brazilian territory and its social and cultural complexity. Thus, using a sample of the speech of 12 informants from Alagoas, of both sexes, distributed in three age groups, this research aimed to analyze 616 occurrences of pronominal subjects of the language spoken in Alagoas, with the intention of observing whether the process of variation and change linguistics already detected by Duarte (1995) in Rio de Janeiro would also be occurring in Alagoas. We adopt a theoretical foundation based on the Theory of Linguistic Variation (LABOV, 1983). The methodology also follows the guidelines of the aforementioned theory. The results revealed that the language spoken in Alagoas is also in a process of change, in relation to the form of realization of the pronominal subjects, since, in the studied corpus, the Alagoas speakers performed 69% of the subject position in its full form and that this process of Linguistic change is affecting Brazilian Portuguese on a very regular basis.

KEYWORDS: Linguistic variation; Subject pronouns; Linguistic change

1 Professora Titular do Centro de Educação e do Programa de Pós-Graduação em Educação, da Universidade Federal de Alagoas-UFAL, maria.cavalcante@cedu.ufal.br.

RESUMO

Diversos estudos demonstram que o Português Brasileiro está em processo de mudança com relação à forma de realização dos sujeitos pronominais. Observa-se uma clara preferência pelas realizações plenas em detrimento dos sujeitos nulos. Entretanto, esses estudos são baseados, em sua maioria, em amostras da fala da Região Sudeste, o que não invalida as pesquisas, mas dá margens a alguns questionamentos, principalmente se levarmos em conta a grande extensão territorial brasileira e sua complexidade social e cultural. Assim, utilizando uma amostra da fala de 12 informantes alagoanos, de ambos os sexos, distribuídos em três faixas etárias, esta pesquisa objetivou analisar 616 ocorrências de sujeitos pronominais da língua falada em Alagoas, com a intenção de observar se o processo de variação e mudança linguística já detectado por Duarte (1995) no Rio de Janeiro também estaria ocorrendo em Alagoas. Adotamos uma fundamentação teórica baseada na Teoria da Variação Linguística (LABOV, 1983). A metodologia também segue as orientações da citada teoria. Os resultados revelaram que a língua falada em Alagoas também está em processo de mudança, em relação à forma de realização dos sujeitos pronominais, pois, no *corpus* estudado, os falantes alagoanos realizaram 69% da posição sujeito na sua forma plena e que esse processo de mudança linguística está afetando o português brasileiro de uma forma bastante regular.

Palavras-Chave: Variação linguística; Pronomes sujeitos; Mudança linguística.

Introduction

Over time, natural human languages have undergone a process of linguistic variation and change; Brazilian Portuguese is no different. Various studies (CAVALCANTE, 2001; DUARTE, 1993, 1995, 1999; ROBERTS, 1993, to name but a few) show that our language is currently going through several changes, among which this study focuses on the ones pertaining to the expression of pronominal subjects, which consequently affect subject position. According to the aforementioned authors, in the first half of the 19th century null pronominal subjects accounted for 80% of subject position expressions, and this did not change significantly until the 1920s, when two simultaneous processes were detected: a marked decrease in null subject usage, and an increase on overt subjects, encompassing both specific reference and arbitrary reference pronominal subjects.

In some Brazilian regions, the current rate of phonetic expression of pronominal subjects is 70%, which means that over the last 100 years there have been changes in how the subject position is filled. Some scholars regard this as a process of linguistic change (DUARTE, 1993, 1995, 1999; ROBERTS, 1993; SHERRE et al. 2015, among others).

However, it is worth mentioning that the process of linguistic variation is not recent, as it Marroquim (1934; 2008 edition), when analyzing the language of the Northeast: Alagoas and Pernambuco, observed that the Brazilian Portuguese spoken in this region was undergoing a process of transformation in inflectional paradigms, both in nominal inflection (absence of the plural mark in nouns and adjectives) and in the verbal inflections, which were presenting a uniformity. That is: the 2nd and 3rd persons had only one form. For Marroquim (2008, p.86), “the absence of nominal inflection will be completed with the abolition of verbal inflections,

affecting mainly persons and tenses, with the role of determining them being reserved almost exclusively for pronouns”. At that time, in the 1930s, the people of Alagoas and Pernambuco still used the pronoun “vos”, however, keeping the verb in the 3rd person singular.

Thus, based on Marroquim (2008) and the research carried out by Duarte (1995), the following questions are raised: is spoken language in Alagoas also undergoing changes in pronominal subject usage? And, if that is the case, are those changes influenced by the same linguistic and extralinguistic factors detected by Duarte (1995) in Rio de Janeiro, and are they happening at a similar rate? Does this possible change affect specific reference and arbitrary reference pronominal subjects at the same rate? Which social factors could be more relevant to this process of linguistic change?

Faced with these questions, this article aims to present the results of a survey carried out from 1997 to 2001, which aimed to analyze 616 occurrences of pronominal subjects in the language spoken in Alagoas, with the intention of observing whether the process of linguistic variation and change, already detected by Duarte (1995) in Rio de Janeiro, would also be occurring in Alagoas. For that, we adopted a theoretical foundation based on the Theory of Linguistic Variation (LABOV, 1983). The methodology also follows the guidelines of the aforementioned theory. In this perspective, the corpus was submitted to the computer program Varbrul (SHERRE 1992/1993).

This study operates on the premise that linguistic change (in this case, involving pronominal subjects) is motivated not just by linguistic factors (phonetic, phonologic, morphologic, syntactic, and semantic ones), but also by extralinguistic ones (gender, age group, education, location, social and historical aspects, etc.). Consistent with this perspective, we subscribe to the Theory of Language Change’s principles and methods; thus, this is a quali-quantitative research, as proposed by Labov (1983).

It is our belief that this type of research can provide valuable contributions to the study of variable linguistic phenomena in Brazil by detecting not only those that seem to be undergoing language changes, but also those that present stable variation.

On one hand, variationist sociolinguistics-based research can contribute to a greater understanding of variable linguistic phenomena, and on the other hand, by electing speech as our object, we are focusing on a modality that was long neglected by cutting-edge linguistic research. According to Moura (1995), the most significant consequence of this neglect is a gap of scientifically-proven studies aiming to shorten the distance between the language spoken by children and the written language taught at schools.

Theoretical and methodological aspects

The theoretical framework underpinning this study is Weinreich, Labov & Herzog’s (1968) Theory of Language Change (see also LABOV 1983; and DUARTE 1993, 1995, 1999a, 1999b), whose method consists of using statistical data to support variable rule linguistic analysis

by assigning different weights to the various factors that might interfere with the probabilities of either form being chosen. The analysis was conducted using Varbrul software.

Variationist tenets

The theory of language change is based on a set of theoretical tenets that broke away from linguistic premises that were firmly established at the time, such as the association between homogeneity and structure. Moreover, many of its postulates introduced great changes to linguistic research, such as the idea that variation is inherent to the linguistic system and that it occurs both in speech communities and in different enunciations produced by a same speaker, according to the social role he or she is playing at the moment. From this standpoint, heterogeneity occurs normally in human languages.

Along with breaking away from Saussurean structuralist postulates, the theory of language change introduced a number of innovations in Linguistics, especially by acknowledging that linguistic variation analyses must consider not only a language's structure, but also social structures, since it is precisely the interaction between social and linguistic aspects that sheds light on matters not previously regarded as linguistic research objects, as this Theory seeks to describe and explain linguistic variation based on the interaction between linguistic and extralinguistic variables, such as a speaker's gender, age group, education, mobility, culture, and history.

Thus, the study of language changes in a variationist perspective implies considering that social factors play a significant role in linguistic variation and linguistic change. Therefore, our investigation of the Brazilian Portuguese pronominal paradigm variation process (and possible linguistic change) requires considering that social factors may be at the source of this process.

According to Weinreich, Labov & Herzog (1968, p.101), "the key to a rational conception of language change – indeed, of language itself – is the possibility of describing orderly differentiation in a language serving a Community". Weinreich, Labov & Herzog's (1968, p.187-188) general principles to capture language change are of the utmost importance to understand the possible ongoing change process involving pronominal subjects in Alagoan spoken language, and thus are quoted below. From this perspective, linguistic change begins when the generalization of a given alternation in a certain subgroup of the speech community comes into play and assumes the character of differentiation.

Linguistic change is not to be identified with random drift proceeding from inherent variation in speech. Linguistic change begins when the generalization of a particular alternation in a given subgroup of the speech community assumes direction and takes on the character of orderly differentiation.

The association between structure and homogeneity is an illusion. Linguistic structure includes the orderly differentiation of speakers and styles through rules which govern variation in the speech community; native command of the language includes the control of such heterogeneous structures.

Not all variability and heterogeneity in language structure involves change; but all change involves variability and heterogeneity.

The generalization of linguistic change throughout linguistic structure is neither uniform nor instantaneous; it involves the covariation of associated changes over substantial periods of time and is reflected in the diffusion of isoglosses over areas of geographical space.

The grammars in which linguistic change occurs are grammars of the speech community. Because the variable structures contained in language are determined by social functions, idiolects do not provide the basis for self-contained or internally consistent grammars.

Linguistic change is transmitted within the community as a whole; it is not confined to discrete steps within the family. Whatever discontinuities are found in linguistic change are the products of specific discontinuities within the community, rather than inevitable products of the generational gap between parent and child.

Linguistic and social factors are closely interrelated in the development of language change. Explanations which are confined to one or the other aspect, no matter how well constructed, will fail to account for the rich body of regularities that can be observed in empirical studies of language behavior (WEINREICH, LABOV & HERZOG, 1968, p. 187-188).

According to Labov (1983), change usually propagates throughout a set of structural environments and is progressively disseminated over more favorable environments. Hence, the matter of social evaluation also bears considerable importance in the study of linguistic change, for it is connected to the degree of attention paid by speakers to the way they speak: certain linguistic change stages can be predicted based on how a speaker evaluates the speech of a given community, because that evaluation allows us to suppose what sort of aspirations that speaker has regarding his/her linguistic behavior, as well as his/her community. Labov (1983, p. 31) posits that the evolution of linguistic change cannot be understood without considering the social life of the community in which it occurs, since language is continuously influenced by social pressures.

As should be clear from above, Weinreich, Labov & Herzogs's (1968, p.187-188) general principles and the five aforementioned matters play a key role in understanding linguistic change phenomena, especially when it comes to data collection. In Labov's (op.263-264) eyes, using a mathematical model or electronic sound devices is not enough, and each data collection step must be executed as properly as possible.

Another, no less important contribution to the study of linguistic variation and change is the development of a suitable methodology to work with linguistic data produced by real speakers in real situations, in keeping with a theoretical framework that does not focus exclusively on linguistic factors, but also considers social ones. This methodology had a formidable impact on speech studies by allowing researchers access to procedures heretofore employed by other sciences, such as the use of mathematical models to aid both the description and the understanding of variable linguistic phenomena.

Regarding the phenomenon already verified by Duarte (1993, 1995) in Rio de Janeiro speech - a process of variation and change in subject position in Brazilian Portuguese involving overt pronouns, we believe that Theory of Language Change tenets can help us detect whether or not this process is occurring in Alagoas at the same rate as in Rio de Janeiro.

Method

This is a quantitative and qualitative research, whose main objective is to compare and analyze, in the light of the Theory of Linguistic Variation, data from the language spoken in state Alagoas, referring to the realization of pronominal subjects, from a synchronous sample of oral language, with data analyzed by Duarte (1995) which also referred to the realizations of pronominal subjects in carioca speech .

The corpus used in this research, which gave origin to this article, is part of the database of the LUAL² Project (Língua Usada em Alagoas) of the Postgraduate Program in Letters and Linguistics at the Federal University of Alagoas. It consists of transcripts of speech recordings of 12 informants from Alagoas, 06 female and 06 male, distributed in three age groups, the first from 15 to 25 years old, the second from 26 to 49 years old and the third from 50 to 80 years old.

The data referring to the 06 informants residing in the interior of the state of Alagoas were collected in the period from 1991 to 1997, by students who participated in the LUAL project; while the recordings of the speeches of the 06 informants residing in Maceió were recorded and also transcribed by the author of this article in 1999 .

The average speech sample duration is of 30 minutes per speaker, recorded in cassette tapes. During a pre-analysis stage, transcripts were screened for the presence of subject pronouns (overt or null) with either defined or arbitrary reference. The next step was selecting 616 sentences with overt and null subject pronouns. The data selection and codification stage were preceded by a pre-analysis in which the linguistic data were submitted to a denaturalization process: rather than using the full transcripts, we selected the excerpts that would comprise our data set. The following step consisted of sentence delimitation; this involved the presence of the factors we intended to analyze and also software parameters - for example, at the time Varbrul did not accept sentence inputs longer than 120 characters.

2 The Lual project was coordinated by Professor Maria Denilda Moura.

Sentences were codified according to independent variable groups, or factor groups. For the purpose of this article, the following linguistic factors were applied: a) reference (defined and arbitrary), verbal inflection and grammatical person; and the following extralinguistic factors: gender, age group, education, and location (whether or not the speaker resided at the state capital). These factors were fed to Varbrul. The software received linguistic information as code and calculated the weight of each independent variable over the dependent binary variable overt pronominal subject vs. null pronominal subject. In addition to each factor's weight or probability, Varbrul also made a statistical selection.

According to Naro (1992), probabilities higher than 0.50 indicate that a factor exerts a positive impact on the variable rule; probabilities under 0.50 indicate a negative impact, and probabilities close to 0.50 indicate that a factor exerts neutral influence over the dependent variable. Naro also explains that Varbrul's analyses involve several levels, as the software compares the probabilistic values assigned to the various factors under study. The first level of analysis is called 'level zero', or rule input; and in each following level, Varbrul calculates probabilities and selects the most relevant ones - the ones most likely to occur.

This statistical treatment of whether or not subject position is filled by an overt pronoun in Alagoan speech is justified when we consider that this phenomenon consists of a choice between two possibilities (overt and null), which may or may not be influenced by several linguistic and extralinguistic factors.

The main goal of the research that originated this study was to analyze Alagoan speech data based on Duarte's (1995, p. 141) research and her claim that "Brazilian Portuguese has lost the characteristic property of pro-drop languages as a result of the impoverishment of its inflectional paradigm". [...] However, "language usage does not reflect that loss as a completed change", which means that Brazilian Portuguese is undergoing a linguistic change process.³

According to Duarte's (1995, p. 48) data, null subject rates from lowest to highest are: 2nd person 20%; 1st person 33%, and 38% for 3rd person. She believes these results suggest that the change process involving subject position began with the 2nd person singular and plural pronouns "tu/vós", that were substituted by variants "você/vocês" respectively, which combine with 3rd person verbal agreements – thus, verbal inflection alone no longer provided enough information to identify grammatical persons, as the same inflection now applied to both the 2nd and the 3rd persons.

The following tables provide an overview of the current pronominal paradigm in Brazilian Portuguese spoken in Rio de Janeiro and in Alagoas, with their respective verbal inflections:

³ Languages characterized by the possibility to fill the subject position either by a null (hidden) or by an overt pronoun.

Table 01. Pronominal and inflectional paradigm of Brazilian Portuguese as spoken in Rio de Janeiro

Person/no.	Pronoun	Paradigm 1	Paradigm 2	Paradigm 3
1 st p. singular	eu	am o	am o	am o
2 nd p. singular	tu	am a s	-	-
	você [var.]	am a s	-	am a
3 rd p. singular	ele/ela	am a	am a	am a
1 st p. plural	nós	am a mos	am a mos	-
	a gente [var.]	-	am a	am a
2 nd p. plural	vós	am a is	-	-
	vocês [var.]	am a m	am a m	am a m
3 rd p. plural	eles/elas	am a m	am a m	am a m

(DUARTE, 1995, p.32)

As seen in Table 01 (1995, p.32), Brazilian Portuguese has evolved from a system comprising six specific verbal inflections (person and number) to a four-inflection system. Moreover, in some regions, the new system has also lost the 2nd person, represented by the “tu” [singular] and “vós” [plural] pronouns. Thus, the pronominal and inflectional paradigm has suffered a rather significant reduction that compromises subject identification to a degree, given that this task can no longer be accomplished exclusively by verbal inflection.

Table 02. Presents the pronominal and inflectional paradigm found in the Alagoan speech sample:

Person/no.	Pronoun	Paradigm 1	Paradigm 2	Paradigm 3
1 st p. singular	eu	am o	am o	am o
2 nd p. singular	tu	-	-	-
	você	am a	am a	am a
3 rd p. singular	ele/ela	am a	am a	am a
1 st p. plural	nós	am a mos	am a mos	am a
	a gente	am a	am a	am a mos
2 nd p. plural	vós	-	-	-
	vocês	am a m	am a m	am a
3 rd p. plural	eles/elas	am a m	am a m	am a

As shown in Table 02, pronominal and inflectional paradigms in spoken language in Alagoas have also suffered a reduction, with the partial loss of the direct 2nd person “tu” [s.] and “vós” [pl.]. The loss of the pronoun “vós” [2nd p. pl.] seems consolidated in Alagoan speech, but the same cannot be said for “tu” [2nd p. s.]; even though it was absent from the sample, its presence can be empirically verified in informal contexts, such as exchanges between father and son, siblings, and teenagers of a same group. However, in these occurrences, verbal inflection usually follows the 3rd p. s. paradigm.

As argued by Duarte (1995, p.33), as a consequence of this reduction in its pronominal and inflectional paradigms, Brazilian Portuguese has come to have “a defective null subject system, in which there is a loss of the overt and null pronominal subject complementarity characteristic of pro-drop languages, as they become interchangeable”. Such interchangeability does not, however, interfere with subject identification and licensing, as subjects are still “licensed and identified according to the same mechanisms used in romance pro-drop languages”. This shows that verbal inflection continues to allow the usage of null pronominal subjects.

In light of this fact, we have chosen to use most criteria employed by Duarte (1995) to organize our data set. Hence, the selected sentences involve time, and a subject position accomplished by an overt or null pronoun (eu, tu, você, ele, ela, a gente, nós, vocês, eles, elas [I, you [s.], you [s. var.], he, she, we [var.], we, you [pl. var.], they [masc.], they [fem.], etc.) of either specific or arbitrary reference.

It is important to stress here that the subject notion should present relationships of dependence and order with the predicate’s central verb; in other words, the analysis must involve the subject pronoun’s agreement with and position in relation to the predicate’s central verb.

A specific reference pronominal subject refers to an identifiable entity that can be mentally recognized by the speaker’s interlocutor; in other words, it corresponds to an image that has been clearly defined by the speaker (CASTILLO, 1987), such as the one in example 01:

(01) - Ah sim tá é, na verdade eu não sô muito de lê jornal não, viu!
[- Oh yes, yes, I’m actually not much of a newspaper reader, you see!] E quando eu leio eu leio a Folha de São Paulo [...] eu prefiro ver televisão, ver as notícias [...] as coisas locais / mais, mas/ é difícil (L3/PNS-5:61) ⁴.
And when I do read I read Folha de São Paulo [...] I prefer to watch tv, to see the news [...] local stuff / but / it’s hard.

In example 01, when the speaker uses the pronoun “eu” [1st p. s.] he is referring to himself (a being that exists in time and space), a referent that can be recognized by his interlocutor and, consequently, this pronoun can be interpreted as a defined reference pronominal subject.

Unlike defined reference pronominal subjects, those with arbitrary reference ones are the those whose referents cannot be recognized by one’s interlocutor as they consist of entities that cannot be identified in time and space, are unrecognizable, generic, or hypothetical – in other words, referents of which the speaker does not produce a clear image; this can result either from an intentional choice, or from lack of knowledge about grammatical representation. Arbitrary reference pronominal subjects are illustrated by the following occurrences of pronoun “você” [2nd p. s. var.] in example 02.

4 Examples are coded as follows: L3, L4 etc. identify the speaker; the capital letters after the slash are the speaker’s initials, followed by the report page number quoted in the example.

(02) -- Hoje Dora, eles vê a religião às vezes como um meio de viver [...] Ø⁵ não estão entendendo o que Deus é para cada ser humano [...], porque você não pode viver sem uma religião você tem ter você tem que obedecer a Deus ou de uma maneira ou de outra ... você tem que viver pra Deus algum minuto de sua vida (L4/GCP-8225)

-- *Nowadays, Dora, they sometimes sees [sing.] religion as a way to live [...] Ø don't understand what God means to each human being [...], because you can't live without a religion you gotta have... you've got to obey God one way or another... you've got to live for God at least a moment of your life*

Occurrences of pronoun “você” in example 02 do not refer to any specific entity, but instead to a generic being for which the speaker does not produce a clear image. In addition to pronoun “você”, we have verified that arbitrary reference subject pronouns also occur in overt pronoun usages involving pronouns “eu” [1st p. s.] and “nós” [1st p. pl.].

Overt pronominal subject vs. null pronominal subject in Alagoan speech

The dependent variable overt vs. null pronominal subject was also controlled for different linguistic factors – reference (defined or arbitrary), verbal inflection, grammatical person, and verb tense, in addition to social ones – speaker gender, age group, education, and location (whether or not they resided in the state capital).

The specific/arbitrary reference variable was included in the study to allow for a twofold investigation: 1) whether the change in specific reference pronominal subjects detected by Duarte (1995) in Rio de Janeiro speech was also affecting arbitrary reference pronominal subjects in Alagoan speech, and 2) whether the conditioning factors that promoted overt pronoun usage in arbitrary reference subjects were the same ones driving the ongoing change observed in specific reference pronominal subjects.

The following tables present pronominal subject occurrence rates in Alagoan and Rio de Janeiro speech, respectively.

Table 03. Pronominal subjects in Alagoas speech

Overt pronominal subject		Null pronominal subject	
Defined reference	Arbitrary reference	Defined reference	Arbitrary reference
69%	69%	31%	31%

Table 04. Pronominal subjects in Rio de Janeiro speech

Overt pronominal subject		Null pronominal subject	
Defined reference	Arbitrary reference	Defined reference	Arbitrary reference
71%	65%	29%	35%

(DUARTE, 1995, p. 47-48)

5 The Ø symbol indicates that the subject position is null, that is, not phonetically filled.

Out of 616 cases selected among our Alagoan speech sample, 514 (83%) are defined reference pronominal subjects. Out of those 514, 353 (69%) are overt subjects while 161 (31%) are null. Arbitrary reference pronominal subjects represent 17% of the speech sample; from this group, 69% are overt subjects while 31% are null.

Duarte's (1995, p.47-48) results from her analysis of Brazilian Portuguese in Rio de Janeiro speech are as follows: 71% specific reference overt pronominal subjects, and 29% null ones. Arbitrary reference pronominal subjects represent 19% of her speech sample; from this group, 65% of occurrences phonetically express subject position, while 35% are null.

A comparison between Tables 03 and 04 reveals some regularity in the numbers. They seem to indicate a degree of uniformity on the ongoing linguistic change process, given that the occurrence rate of specific reference pronominal subject is 69% in the Alagoan speech data, and 71% in Duarte's data. Null pronominal subject occurrence rates are equally similar: 31% in the Alagoan data and 29% in the Rio de Janeiro data. This relative symmetry also applies to the occurrence rates of arbitrary reference pronominal subjects: 69% and 65% for overt pronominal subjects, and 31% and 35% for null ones.

Such data regularity seems to indicate that, although those studies were conducted in different contexts (Alagoas and Rio de Janeiro), there are many similarities in Brazilian Portuguese spoken across different regions - and those similarities appear to be more evident in terms of syntax.

Furthermore, results show that the change process affecting specific reference pronominal subjects is also impacting arbitrary reference ones. Indeed, when those types of pronominal subjects are analyzed separately, controlling the dependent variable for the chosen linguistic and extralinguistic factors, results indicate which factors have the greatest impact on overt pronominal subject occurrences in both specific reference and arbitrary reference constructions.

Regarding specific reference occurrences, Varbrul selected **verbal inflection** and **grammatical person** as relevant factors of the ongoing linguistic change involving pronominal subjects in Alagoan speech. Among social factors, the location variable stands out as the one bearing the greatest weight: .59. The software detected the same linguistic factors as most significant for arbitrary reference occurrences, while **gender** and **location** were the social factors with the greatest impact on this set of occurrences.

Linguistic data analysis⁶

From a linguistic standpoint, the independent variable verbal inflection ranked first in relevance to overt pronominal subject expression both in the specific reference occurrence set and in the arbitrary reference one.

Table 05 below presents the results of controlling the dependent variable verbal inflection for specific reference pronominal subjects in the Alagoan speech data set.

Table 05. Specific reference pronominal subject vs. verbal inflection

Inflection/person	-o, -i 1 st p. s.	-mos 1 st p. pl.	-m 3 rd p. pl.	-0(zero) 1 st , 2 nd , and 3 rd p. s./pl.
Occurrences	294	10	35	170
Overt subject	189	7	21	133
Null subject	105	3	14	37
Weight	.44	.51	.40	.67
Significance level = .000				

The mathematical values depicted in Table 05 show that although affixes **-o** and **-i** (1st p. s.) are used on many overt pronominal subject occurrences, the .44 weight indicates that this factor does not positively impact them. This low weight value seems to suggest that 1st person singular null pronominal subject occurrences will last for a long time in Alagoan speech.

Unlike the 1st person singular morpheme **-o (i)**, which has a high occurrence rate (294 cases, 189 overt and 105 null), the 1st person plural **-mos** affix is used in only 10 occurrences (7 overt subjects and 3 null subjects); in other words, while the **-o (i)** affix is present in nearly half the sentences in our sample, the **-mos** affix is present in less than 2% of them, indicating that it will not integrate spoken Brazilian Portuguese for long.

As for the **2nd and 3rd** persons plural **-m** affix, in our data set it only occurs with 3rd p. pl. expressions. Notably, although 60% of the subjects in this group are overt, the .40 weight indicates that the **-m** verbal inflection does not positively impact overt pronominal subject occurrences. This does not mean that there is no change in progress, merely that the process involving this particular affix appears to be taking place at a slower pace.

These **-m** affix results are quite compatible with Duarte's (1995) and also with Monteiro's (1994), whose study shows a 60% rate of overt third person plural subjects.

⁶ We would like to stress the fact that this study reports part of the findings of a more comprehensive research in which overt specific reference pronominal subjects were controlled for 13 factors, from which 7 were found to impact overt subject occurrence. However, in this paper we only analyze the variables selected by Varbrul software as the most significant, according to their weights.

Whilst the three aforementioned affixes had frequency rates above 60% but their insignificant weights indicated a negative impact on speeding up the linguistic change under analysis, the **-0 (zero)** affix, which can express all grammatical persons, presents not only the highest frequency rate (78%) in overt pronominal subject occurrences, but also the highest weight: .67. This leaves no doubt about its influence in the linguistic change process affecting pronominal subjects in Alagoan speech.

According to Duarte (1995), the high frequency of the **-0 (zero)** affix shows an ongoing impoverishment of the Brazilian Portuguese inflectional paradigm. She goes so far as to argue that this phenomenon can contribute to make null pronominal subjects progressively rarer over time, or even marginal. Examples 03, 04, 05, 06, 07, and 08 below illustrate **-0** affix occurrences.

- (03) Exato eu fazia **(-0)** curso até por correspondência (L2/JVS-12:328)
Exactly, I even took (-0) correspondence courses.
- (04) Olha – você é **(-0)** daqui (L6/AFM-22:552)
Look – you're (-0) from here
- (05) Ela é já hoje – já mora **(-0)** com um home (L8/EGM-33:842)
Nowadays she is – she already lives (-0) with a man
- (06) Como professor a gente já pega **(-0)** muita experiência (L4/GCP-08:211)
As teachers we gain (-0) a lot of experience
- (07) Nós comprou **(-0)** geladeira (L11/NEC-29:58)
We bought (-0) [3rd p. s. inflection] a fridge
- (08) Eles me atende **(-0)** bem né? (L3/FJS-02:23)
They serve (-0) [3rd p. s. inflection] me well, right?

As examples 03 through 08 make clear, the **-0** affix is being used to express all grammatical persons in the Alagoan speech sample. The sentences in examples 07 and 08 were chosen to demonstrate that the **-0** affix does not exclusively accompany pronouns with singular verbal agreements, but also plural-marked pronouns (we, they). The examples also stress the fact that the **-0** inflection ranks among the factors with the greatest positive impact on the linguistic change process affecting subject position expression in the Alagoan speech data set, hence Varbrul selecting it as the most significant variable.

According to Roberts (1993, p. 415), French was also a null subject language once, but the process that eventually led to the loss of that property lasted for 150 years. During that period, French had a defective null subject system that allowed for null pronominal subject expression only in certain syntactic contexts and involving certain grammatical persons.

Aside from **verbal inflection**, overt and null pronominal subjects were also controlled

for the **grammatical person** variable in order to gain a better understanding of the relationship between this variable and verbal inflection; we believed that analyzing the grammatical person variable could provide further elements to show how the relationship between verbal inflection and grammatical persons affected the linguistic change process under analysis (overt vs. null pronominal subject expression). Table 06 below presents the results of controlling specific reference pronominal subjects for **grammatical persons**.

Table 06. Defined reference pronominal subject vs. grammatical persons in Alagoan speech

Person	eu [1st p. s.]	você [2nd p. s. var.]	ele/ela [3rd p. s. masc./fem.]	nós [1st p. pl.]	a gente [1st p. pl. var.]	eles/elas [3rd p. pl. masc./fem.]
Occurrences	355	21	60	10	28	35
Overt subject	239	11	40	7	27	26
Null subject	116	5	20	3	1	14
Weight	.47	.48	.48	.50	.92	.44
Significance level = .000						

As shown in Table 06, overt subjects are preferred across all grammatical persons in Alagoan speech. The 1st person has the highest occurrence rate; this demonstrates its relationship with 1st p. s. verbal inflection **-o (-i)**, which was also the most productive. 1st p. s. pronoun “eu” has the highest occurrence rate in our sample, despite being associated to an exclusive affix that could, in principle, convey grammatical person through verbal inflection and thus favor null subjects. However, its .47 weight indicates that that variable has no significant impact on the linguistic change investigated in this study.

Paradigmatic 1st p. pl. pronoun “nós” also bears neutral weight on the pronominal subject change process; data shows that it is seldom used. Speakers heavily favor the innovative pronoun “a gente” [1st p. pl. var.], whose .92 weight makes it indisputably relevant to the linguistic change process.

The preference for “a gente” over “nós” was also verified by Vianna and Lopes (2015, p. 109) in their analysis of linguistic variation involving pronouns “nós” and “a gente” in Brazilian Portuguese. According to the authors, this phenomenon “may be regarded as a linguistic change process in which the innovative option increasingly gains terrain over the older one”.

Lopes (1992) also found a high rate of overt subject constructions for pronouns “nós” and “a gente” in her analysis of the occurrence of those pronouns in Standard Brazilian Portuguese. In her sample, the rate of overt subjects involving “nós” and “a gente” was 67% and 81%, respectively. Example 09 below illustrates specific reference cases using the pronoun “nós”.

(09) a vida como é ... como se apresenta a nós...eu fui aluno dele... depois nós fizemos concurso pra...pra escrivão de Coletoria Federal...nós fomos nomeados e ele passou a ser meu colega ... e eu colega dele (L6/AFM-24:608). *such is life... so it happens... I was his student... then later we both applied for... for Federal Tax Clerk positions... we were accepted, and he became my coworker... and I, his coworker.*

Both instances of pronoun “nós” in example 09 show that the speaker prefers to fill the subject position with a pronoun even when using the exclusive 1st p. pl. affix **-mos**. Interestingly, the sentence in example 09 was produced by a 72-year-old male informant with a middle-school education; this suggests that the ongoing linguistic change process also affects senior speakers.

2nd p. s. pronoun “você” originated from honorific “vossa mercê” then proceeded to evolve into the sole 2nd p. pronoun in most Brazilian Portuguese dialects, especially given the partial disappearance of 2nd p. s. pronoun “tu” and the definitive, irreversible loss of 2nd p. pl. pronoun “vós”. Scherre’s (2015, p. 171) analysis of pronominal variation involving “tu” and “você” verified that there are already reduced variants of the second – “cê” and “ocê”. While the latter is characteristic of certain regions/social contexts, “cê” is used in practically all Brazilian speech communities studied by the author across the country.

According to Duarte (1995, p. 48), 2nd p. pronoun “você” has the lowest null subject rate, around 11%, and correspondingly, the highest rate of filled subject positions.

That author believes that the low occurrence rate of 2nd person null subjects in Rio de Janeiro speech indicates that it “triggered off the linguistic change process – due to its substitution by “você” [...], which combines with 3rd p. verbal agreement – while also being the grammatical person to soonest present the effects of that change”.

Whilst there were no occurrences of 2nd p. pl. pronoun “vocês” in our sample, our empirical experience allows us to verify that it is indeed the sole pronoun used in Alagoan speech to indicate 2nd person plural, since “vós” is no longer in use.

The null subject rate for 3rd p. s. masculine and feminine pronouns “ele/ela” was also low: 37%, with the corresponding high rate of overt subjects being 67%; this suggests that 3rd p. s. pronominal subjects follow the same trend observed in pronouns “eu”, “você”, and “a gente”. There were few occurrences of 3rd p. pl. constructions in our sample. Despite the prevalence of overt subjects, this grammatical person has the highest null subject rate.

Social variables

The dependent variable overt vs. null pronominal subjects was also controlled for the following social factors: **gender**, **age group**, **education**, and **location** (whether or not the speaker resided at the State capital). Varbrul has selected **gender** and **location** as the variables with the most relevant impact on overt subject occurrence.

Controlling the dependent variable for extralinguistic factors is of the utmost importance; several variationist sociolinguistic studies prove that aside from playing a significant part in linguistic variation and change, social factors can also be the ones to trigger off a change process (LABOV, 1983; MOLLICA, 1992; SHERRE, 1992, 1996; NARO, 1992; PAIVA, 1992; SILVA-CORVALAN, 1989, among others).

Hence, based on Duarte (1995), Paiva (1992), and Votre (1992), we posit that female Alagoan speakers tend to use overt pronominal subjects significantly more often than male speakers.

Women were found to use overt subjects at a .55 probability rate and 70% frequency rate, against men's respective rates of .43 and 59%.

We chose to control for location with the aim of investigating whether there was a quantitative difference in overt pronominal subject occurrence rates between speakers who reside at the capital city and elsewhere in the State.

Varbrul results show a considerable difference: the overt subject rates of non-capital speakers are 74% frequency and .59 probability, against capital speakers' 58% frequency and .41 probability.

It is important to stress that we did not set out to conduct an in-depth, exhaustive analysis of all possible causes of the predominance of overt pronominal subjects in the speech of non-capital informants; that task would exceed the scope of this study, since different sorts of factors (linguistic, social, psychological) likely contribute to have non-capital folk spearhead the linguistic change process.

Conclusion

This paper reports part of the findings of a more comprehensive research that analyzed 616 pronominal subject occurrences in Alagoan speech in order to verify whether the process of linguistic variation and change detected by Duarte (1995) in Rio de Janeiro was also happening in Alagoas. To that end, the theoretical framework underpinning this study is Labov's (1983) Theory of Language Variation, which also informs the research methodology. The corpus was analyzed using Varbrul software (SHERRE 1992/1993) and the results show that Alagoan speech is also undergoing a change in pronominal subjects: the overt subject rate found in our sample was 69%.

Results also show that most of the conditioning factors influencing linguistic change in Alagoas are the same as those found by Duarte (1995) in Rio de Janeiro. Moreover, we have also verified that the change process affects both specific reference and arbitrary reference pronouns: they have the same overt pronominal subject rate of 69%.

At the morphologic level, verbal inflection was selected by Varbrul as the variable with the most significant impact on linguistic change. The most significant factors at that level are inflection “0” (zero) and pronoun “a gente” [1st p. pl. var.], both of which present the highest frequency and probability rates of overt pronominal subjects in the analyzed Alagoan speech sample. In addition, we have verified that verbal inflection “0” is combined by speakers with all grammatical persons; hence, this factor contributes unequivocally and significantly not only to disseminate the usage of overt pronominal subjects, but also to implement a change in the Brazilian Portuguese pronominal paradigm.

Finally, gender and location were selected by Varbrul as the extralinguistic variables with the most relevant impact on overt pronominal subject usage. This indicates that the linguistic change has progressed further off-capital, and that women are leading this implementation.

References

CAVALCANTE, M. A. da S. *O sujeito pronominal em Alagoas e no Rio de Janeiro: um caso de mudança em progresso*. 2001. Maceió: UFAL. Tese de Doutorado. Programa de Pós-Graduação em Letras e Linguística da Universidade Federal de Alagoas. 2001.

CASTILLO, A. T. de. *A ordem do sujeito nominal do português culto falado em São Paulo*. São Paulo: Unicamp, 1987.

DUARTE, M. E. L. Do pronome nulo ao pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I; KATO, M A. (org.). *Português Brasileiro: uma viagem diacrônica*. Campinas: Ed. Da Unicamp, 1993, p.107 – 128.

DUARTE, M. E. L. *A perda do princípio evite pronome no Português Brasileiro*. 1995. Tese de Doutorado – Universidade Estadual de Campinas. 1995.

DUARTE, M. E. L. *O sujeito nulo no português europeu e no português brasileiro*. Londrina: Editora da Universidade Estadual de Londrina, 1999a.

DUARTE, M. E. L. A Sociolinguística paramétrica: perspectivas. In: HORA, D; CHRISTIANO, E. *Estudos linguísticos: realidade brasileira*. João Pessoa: Ed: Ideia, 1999b.p. 107-114.

LABOV, W. *Modelos sociolinguísticos*. Madrid: Ediciones Cátedra, 1983. Traducion de José Miguel Herreras. Título original: *Sociolinguistic Patterns*. Philadelphia, University of Pennsylvania Press, 1972.

LOPES, C. R. *Nós e a gente no português falado culto*. Rio de Janeiro, 1992. Dissertação de Mestrado – Faculdade de Letras – Universidade Federal do Rio de Janeiro. 1992.

MARROQUIM, M. *A língua do Nordeste: Alagoas e Pernambuco*. 4 ed. Maceió: EDUFAL, 2008.

MOLLICA, M. C. (org). Introdução à sociolinguística variacionista. *Cadernos didáticos*. Rio de Janeiro: Ed. Da UFRJ, 1992.

- MOLLICA, M. C. *Como o brasileiro fala, percebe e avalia alguns padrões linguísticos*. Rio de Janeiro: Editora Flores Verbais, 1995.
- MONTEIRO, J.M. *Pronomes pessoais: subsídios para uma gramática do Português do Brasil*. Fortaleza; Edições UFC, 1994.
- MOURA, Maria Denilda. *Variação e ensino*. Maceió: EDUFAL, 1997.
- MOURA, Maria Denilda. *Os múltiplos usos da Língua*. Maceió: EDUFAL, 1999.
- NARO, A. J. Modelos matemáticos e o tratamento estatístico. In: MOLLICA, M. C. Introdução à sociolinguística variacionista. *Cadernos didáticos*. Rio de Janeiro: Ed. da UFRJ, 1992.
- PAIVA, M. da C. Sexo. In: MOLLICA, M. C. Introdução à sociolinguística variacionista. *Cadernos didáticos*. Rio de Janeiro: Ed. da UFRJ, 1992.
- ROBERTS. I. O português brasileiro no contexto das línguas românicas. In: ROBERTS. I; KATO, M.A. (org.). *Português brasileiro: uma viagem diacrônica*. Campinas: Ed. Da Unicamp, 1993.
- SCHERRE, M. M. P. Levantamento, codificação, digitação e quantificação dos dados. In: MOLLICA, M. C. Introdução à sociolinguística variacionista. *Cadernos didáticos*. Rio de Janeiro: Ed. da UFRJ, 1992.
- SCHERRE, M. M. P. *Introdução ao pacote Varbrul para microcomputadores*. UFRJ/UNB, 1992/1993. Mimeo.
- SCHERRE, M. M. P. Pressupostos teóricos e suporte quantitativo. In: Oliveira e Silva, G.; SCHERRE, M.M.P. (orgs) *Padrões sociolinguísticos: análises de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Ed. Tempo Brasileiro- UFRJ, 1996.
- SCHERRE, M. M. P. et.al. Variação dos pronomes “tu” e “você”. In: MARTINS, M.A; ABRAÇADO, J. (org.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015.
- SILVA-CORVALÁN, B. *Sociolinguística – teoria y análisis*. Madrid: Editorial Allambra, 1989, p. 151-199.
- VIANNA, J. S; LOPES, C.R. dos S. Variação dos pronomes “nós e a gente”. In: MARTINS, M.A; ABRAÇADO, J. (org.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015.
- VOTRE, S. Escolaridade. In: MOLLICA, M. C. Introdução à sociolinguística variacionista. *Cadernos didáticos*. Rio de Janeiro: Ed. da UFRJ, 1992.
- WEINREICH; LABOV; HERZOG. Empirical foundations for a theory of language e change. In: LEHMAN & MAIKIEL (ed.). *Directions for historical linguistics*. Austin & London: University of Texas Press, 1968.



A VARIAÇÃO NÓS E A GENTE EM FORTALEZA NA SEGUNDA DÉCADA DOS ANOS 2000: FATORES LINGUÍSTICOS

VARIATION OF THE SUBJECTIVE PRONOUNS NÓS AND A GENTE IN THE SECOND DECADE OF THE 2000: LINGUISTIC FACTORS

Maylle Lima Freitas¹

Lorena da Silva Rodrigues²

Hugo Leonardo Gomes dos Santos³

RESUMO

Esta pesquisa se atém ao estudo da variação dos pronomes sujeito *nós* e *a gente* por falantes de Fortaleza-CE com ensino superior na segunda década dos anos 2000. O estudo baseia-se nos pressupostos teórico-metodológicos da Teoria de Variação e Mudança (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]), em vias de analisar o encaixamento da forma inovadora *a gente* no sistema linguístico na comunidade de fala de Fortaleza-CE e a provável mudança linguística em progresso. Partiu-se da análise de 18 entrevistas do Projeto Descrição do Português Oral Culto de Fortaleza (PORCUFORT) - fase II, na modalidade de registro DID (Diálogo Informante e Documentador). Dos fatores linguísticos, analisou-se o tempo e tipo de paradigma do verbo, a referencialidade do verbo (genérica ou específica) e o tipo de verbo. Utilizou-se o *software* RStudio para o tratamento dos dados, elaboração dos testes estatísticos e criação de modelos de regressão logística para análise dos dados linguísticos. Constatou-se a predominância da forma *a gente* com 82,22% (N = 985) dos dados. A forma *a gente* mostrou-se favorecida pela referência genérica, tempo verbal presente de forma idêntica ao pretérito perfeito do indicativo e pretérito imperfeito. A forma *nós* parece resistir no tempo verbal pretérito perfeito idêntico ao presente (ontem comemos/ hoje comemos) e referente mais específico. O tipo de verbo não se mostrou um fator condicionador a variação linguística nos modelos de regressão logística, ainda que tenha sido observado uma menor força da forma *a gente* com verbos de estado. Por fim, percebemos que a forma *a gente* está altamente implementada no sistema linguístico, em um provável processo de mudança em curso, enquanto a forma *nós* parece ainda resistir como maneira de evitar ambiguidade semântico temporal em caso de formas idênticas no presente e pretérito perfeito do indicativo e enfatizar a referência específica do verbo.

PALAVRAS-CHAVE: Teoria de Variação e Mudança; Variação nós e a gente; Comunidade de fala de Fortaleza-CE; Descrição do Português Brasileiro.

1 Mestranda em Linguística, Universidade Federal do Ceará (Programa de Pós-Graduação em Linguística). E-mail: mayllemalimafreitas@gmail.com.

2 Doutora em Linguística, Universidade Federal do Ceará (Programa de Pós-Graduação em Linguística). E-mail: lorena.rodrigues@letras.ufc.br.

3 Doutorando em Linguística, Universidade Federal do Ceará (Programa de Pós-graduação em Linguística). E-mail: lprof.hugoleo13@gmail.com.

ABSTRACT

This research aims to analyze the subject pronouns *nós* e *a gente* by Fortaleza graduate speakers in the second decade of 2000. The study is based on the Variation and Change theory (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) aiming to investigate the fitting on Fortaleza's speech community linguistic system of the innovate variant *a gente* and a probable linguistic change in progress. The database was 18 interviews from the project Graduate's Oral Portuguese of Fortaleza - phase II (PORCUFORT), on an interviewer and informant register. From linguistic factors, were analyzed: verbal tense and paradigm, verb 's reference (generic or specific) and verb type. Data treatment, statistical analyses and logistic regression models were conducted on Rstudio software. The innovative form *a gente* is predominant on our data, representing 82.22% (N = 985) of the sample, present tense, regarding the ambiguous forms, identical to past perfect tense form, imperfect and generic reference has favored this variant. The conservative form *nós* seems to rely on past perfect verbal tense and specific reference. Verb type was not a relevant factor on logistic regression models, however it was observed a minor presence of *a gente* variant with *to be* verb type. In conclusion, it was noticed that the innovative variant *a gente* is implemented on Fortaleza's linguistic system, on a potential linguistic change in progress, meanwhile the variant *nós* seems to emphasize the specific referent and resist as a resource to avoid semantic-temporal ambiguity due to Portuguese's identical present and past perfect verbal forms.

KEYWORDS: Linguistic Variation and Change; *Nós* and *A gente* variation; Fortaleza's speech community; Brazilian Portuguese studies.

Introdução

A Teoria de Variação e Mudança Linguística (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) tem como pressuposto coexistência de formas linguísticas de mesmo significado referencial. A expressão variável de primeira pessoa do plural, a variação das formas pronominais *nós* e *a gente*, é um fenômeno amplamente estudado no Português Brasileiro (PB). Considerando investigações de cunho variacionista, existem pesquisas com dados de fala do Projeto Norma Linguística Urbana Culta (NURC), desde a década de 1970, que indicam a implementação e a alta frequência de uso da forma inovadora *a gente* em diversas variedades do PB (ZILLES, 2007). Este estudo tem como objetivo investigar o fenômeno em tela na fala de pessoas com Ensino Superior, nativas e residentes de Fortaleza, capital do estado do Ceará, na segunda década dos anos 2000, a partir de dados de fala oriundos do Projeto Português Oral Culto de Fortaleza - Porcufort (Fase II), em vias de contribuir com a descrição dessa variação e expansão do mapa de estudos sociolinguísticos do fenômeno.

Consideramos uma variável linguística morfossintática binária, composta da forma *nós* acompanhada de verbo com desinência de primeira pessoa do plural *-mos* (1) e a forma *a gente* com concordância singular, desinência zero (2), na função de sujeito da oração.

(1) teve um dia que **nós** fomos pra Lençóis:... é LONge... **nós** fomos pra Lençóis:... é LONge... **nós** saímo de lá dez horas, aí **nós** fomos logo atrás das pousada. (INQ.24).

(2) eu acho que é/ as vezes **a gente** se acomoda de ficar dentro né com a família... mas vezes **a gente** enjoa DEmais ta só a com a cara do outro né? mas **a gente precisa**... (INQ. 13).

Para esta análise, nos ateremos nas variáveis linguísticas: tempo e tipo de paradigma do verbo, tipo de verbo, referência do verbo (genérica ou específica) em entrevista sociolinguística de registro Diálogo Informante e Documentador (DID).

O tratamento estatístico dos dados foi realizado por meio da linguagem de programação R, por meio do *software* RStudio - R Core Team (2021), para análise descritiva, frequências e proporções, como também para análise de estatística inferencial, por meio de testes estatísticos de qui-quadrado e modelos de regressão logística, buscando confiabilidade dos resultados como projeção de uso da comunidade estudada.

Este artigo está estruturado da seguinte maneira: no tópico “As formas *nós* e *a gente*: um panorama da variação”, apresentaremos o fenômeno estudado e o estado da arte. Nos pressupostos teóricos “Teoria da variação e Mudança” nos atermos a teoria de base dos estudos variacionistas, na seção seguinte nos debruçamos sobre a metodologia, seguida pela análise dos resultados, e, por fim, as considerações finais, nossa síntese e reflexões sobre a investigação, encerramos com as referências dos estudos.

As formas *nós* e *a gente*: um panorama da variação

A forma *a gente*, apesar de ser a forma inovadora de primeira pessoa de plural, já se encontra implementada no sistema linguístico do Português Brasileiro há séculos, existindo registros da forma na escrita desde o século XIII (LOPES, 2007). Pressupõe-se que, devido a um processo de gramaticalização, o artigo “a” combinado ao substantivo “gente”, devido ao traço semântico de coletividade da expressão, relaciona-se a noção de “eu-ampliado” e passou a exprimir o valor de primeira pessoa do plural e coocorrer com o pronome previamente no sistema linguístico do Português Brasileiro, o *nós*, ainda que tenha mantido sua marca verbal de terceira pessoa do singular na concordância verbal.

Existem diversos estudos que contemplam essa variação em diversas comunidades de fala por todo o Brasil, a exemplo de Omena (1996, 2003), Zilles (2007), Maia (2009), Vianna e Lopes (2015), Freitag (2016), Scherre, Yacovenco e Naro (2018a, 2018b), para expor apenas algumas investigações que estudamos acerca do fenômeno. Sendo a forma predominante em grande parte das comunidades de fala estudadas, com destaque a Zilles (2007) que indica a predominância de mais de 80% da forma *a gente* em dados do projeto NURC. Percebe-se, como tendência geral do PB, a tendência à realização das formas pronominais seguidas dos verbos com concordância padrão, sendo a forma *nós* sem *-mos* (*nós estava*) restrita a comunidades de fala e/ou grupos sociais específicos. A forma *a gente* com *-mos* é rara, existindo possivelmente devido a casos de hipercorreção gramatical, sendo normalmente desconsiderada de estudos devido a sua baixíssima aparição em *corpora*.

Tratando-se da variedade de Fortaleza-CE, existem os estudos de Araújo (2018), com dados de falantes escolarizados até o ensino médio na década de 2000 - *corpus* NORPOFOR;

Silva (2020), com dados de falantes até ensino médio na década de 1980 - *corpus* DSC; Carvalho, Freitas e Favacho (2020), com dados de falantes com ensino superior na década de 1990 - *corpus* Porcufort e Freitas, Rodrigues e Santos (2021), com dados em tempo real com dados de falantes com ensino superior nas décadas de 1990 e 2010 - *corpus* Porcufort fases I e II. Sendo a variação da comunidade predominantemente dicotômica entre as formas pronominais com concordância padrão, a exceção de idiosincrasias. Percebemos em Fortaleza-CE um avanço mais lento da forma *a gente*, em comparação a outras capitais no mesmo período de tempo, em que Carvalho, Freitas e Favacho (2020) apontam um total de 62% da variante.

Atesta-se no Português Brasileiro uma tendência de uso a favor da forma inovadora *a gente* em diversos estratos sociais. Zilles (2007), Freitag (2016), Freitas e Carvalho (2020) e Freitas, Favacho e Carvalho (2022), indicam em seus estudos que a forma inovadora *a gente* não parece sofrer estigma, um fator positivo para sua expansão e que favorecia os possíveis processos de mudança em curso atestados em estudos aqui citados. Ainda que ambas as formas coexistam, devemos considerar fatores condicionantes ao uso das formas variantes, a exemplos de fatores sociais como sexo, escolaridade, faixa etária, e fatores linguísticos como a referência do verbo, saliência fônica, tempo e tipo de paradigma *etc.*

Tempo e tipo de paradigma do verbo

O tempo e paradigma do verbo é uma variável linguística que tem alçado grande importância nos estudos da variação *nós* e *a gente* no português Brasileiro, sendo o grupo criado de acordo com Scherre, Yacovenco e Naro (2018a), baseados em considerações de Naro, Görski e Fernandes (1999) sobre saliência fônica. A primeira questão de interesse é que, considerando a desinência canônica de primeira pessoa do plural (-*mos*), verbos regulares apresentam a tendência de apresentar formas idênticas no pretérito perfeito e no presente do indicativo, por exemplo: *nós passeamos com o cachorro hoje/ nós passeamos com o cachorro ontem*. Com isso, estudos variacionistas atestam a tendência do uso da forma *a gente*, com desinência de terceira pessoa do singular, como maneira de evitar ambiguidades semântico-temporais (*a gente passeia com o cachorro hoje/ a gente passeou com o cachorro ontem*).

Considerando a saliência fônica, Naro, Görski e Fernandes (1999) discorrem sobre a oposição singular/plural e a diferenciação de matéria fônica como fatores importantes à variação de primeira pessoa do plural. Tendo como base a desinência padrão de primeira pessoa do plural, existem maiores ou menores diferenças de matéria fônica quando comparado a desinência de terceira pessoa do singular. Ou seja, quando comparamos a forma *comíamos/comia* temos a manutenção da sílaba tônica e, apenas, a perda da desinência -*mos*, o que configura uma baixa saliência fônica. O mesmo não acontece na oposição *somos/é*, em que existe uma completa mudança do material fônico devido a irregularidade verbal, o que configura uma alta saliência fônica. Alguns tempos verbais apresentam tendência a menor saliência fônica, como

os do indicativo, e outros a maior saliência fônica, como por exemplo o pretérito perfeito do indicativo. Com isso, obtemos 5 níveis para a variável tempo e tipo de paradigma do verbo, descritas abaixo:

1. Presente com forma igual ao pretérito: casos em que a desinência canônica de primeira pessoa do plural é idêntica no presente do indicativo e pretérito perfeito do indicativo, gerando potencial ambiguidade semântico-temporal.

(3) bem simples **nós produzimos** fios, mas pra produzir fios só algodão, poliéster... tem vários tipos de fibra (INQ. 03)

2. Pretérito perfeito igual ao presente: casos em que a desinência canônica de pretérito perfeito é idêntica ao presente do indicativo, gerando potencial ambiguidade semântico-temporal.

(4) viagem excelente, de uma vez só **nós conhecemos** cinco países né (INQ. 50)

3. Presente de forma diferente do pretérito: comum em verbos irregulares, são casos em que não existe ambiguidade entre as desinências canônicas de primeira pessoa do plural do presente ao pretérito perfeito.

(5) Ah, também **nós não temos** um parque industrial aqui (INQ. 3)

4. Pretérito perfeito diferente do presente: comum em verbos irregulares, são casos em que não há ambiguidade entre as desinências canônicas de primeira pessoa do plural do pretérito perfeito ao presente, apresentam tendência a altos níveis de saliência fônica, devido a maior diferenciação de material fônico na oposição singular/plural.

(6) Quando minha terceira irmã nasceu **nós sempre fomos** muito próximas (INQ. 7)

5. Pretérito imperfeito: casos de forma verbal proparoxítona, de tendência a baixa saliência fônica e manutenção da sílaba tônica na oposição singular/plural.

(7) eu e mais oito irmãos, **nós tínhamos** muita liberdade, certo? (INQ. 13)

Com isso, o grupo condensa nesses cinco níveis tanto o tempo verbal, quanto a existência, ou não existência, de ambiguidade, como também atesta os grupos de maior saliência fônica, sendo uma variável complexa e de grande potência explicativa à variação entre *nós* e *a gente*.

Referencialidade do verbo

A referencialidade do verbo é um grupo semântico que trata da referência genérica ou específica do pronome. A partir de Lucchesi (2009) e Carvalho, Freitas e Favacho (2020), consideramos estes dois níveis da seguinte maneira:

Referência específica: os referentes podem ser determinados de maneira explícita ou pelo contexto (8), em que podemos recuperar os sujeitos da oração. O nível também inclui casos em que o falante utiliza a forma de primeira pessoa do plural para referir-se apenas a si mesmo (9), tal como o plural de modéstia.

(8) minha relação com as **minhas irmãs** no início, quando nós éramos menores, nós brigávamos muito (INQ. 7)

(9) a gente vive o luxo... o prédio tua casa... né... é o que a gente dá valor... **eu** sou assim... (INQ. 60)

Referência genérica: trata-se da maior indeterminação do sujeito, abrange a delimitação circunscrita, que se refere a um grupo grande e indeterminado de pessoas sob o mesmo égide (10), podendo ser funcionários de uma mesma empresa ou torcedores de um mesmo time, por exemplo. Como também considera a indeterminação universal (11), em que pode-se referir a toda a humanidade, ou uma grande parcela dela.

(10) dinâmica de leitura, roda de conversa de leitura, na/no **CEJA** a gente não consegue fazer isso porque os alunos eles são muito/eles tem quase que postura acadêmica (INQ. 32)

(11) **todos** nós somos políticos... eu acho... eu penso assim... todos nós somos políticos... (porque) num sei... às vezes eu vejo **o ser humano** dessa face... (INQ. 60)

Com isso, codificamos os níveis da variável a partir das categorias definidas a partir dos estudos de base, tendo como ponto de decisão o traço semântico [+ determinado] para referência específica e o traço semântico [- determinado] para referência genérica.

Tipo de verbo

O tipo de verbo é uma variável que analisa o valor semântico do verbo, categorizado em cinco tipos: verbo ter, verbo que indica estado, verbo epistêmico (conhecimento), verbo *dicendi* (do falar/dizer), verbo que indica ação. Essa variável foi controlada com base em Araújo (2018) cuja pesquisa selecionou o tipo de verbo como fator significativo estatisticamente para o condicionamento do fenômeno. Os exemplos a seguir ilustram cada um dos tipos de verbo analisados.

Verbo ter

(12) uma coisa que me dói muito... é a gente **ter** o teatro José de Alencar e ninguém conseguir entrar no teatro... (INQ. 60).

Verbo que indica estado

(13) hoje em dia **nós somos** muito amigos... mas foi meio complicado (INQ. 37).

Verbo epistêmico

(14) se *a gente* vai **analisar** a mulher ... no Império Romano por exemplo ... ou a Mulher ... na Grécia por exemplo ... (INQ. 48).

Verbo *dicendi*

(15) - não, é rigoroso por que:...éh:... *a gente* **conversa** com o chefe (INQ. 34).

Verbo que indica ação

(16) só que eu sou ansiosa enquanto ele não chegou eu não me sosseguei, mas pronto quando ele chegou *nós* **botamos** as coisas no carro e fomos (INQ. 24).

As variáveis linguísticas apresentadas serão, nesta pesquisa, interpretadas à luz da Teoria da Variação e da Mudança Linguística, cujos principais conceitos serão apresentados na seção a seguir.

Teoria de Variação e Mudança Linguística

A Teoria da variação e Mudança Linguística entende a língua como um sistema social e heterogêneo. Tal heterogeneidade é passível de ser descrita e sistematizada, a partir de estudos que buscam investigar os fatores que levam à variação linguística. Entende-se por variação “o processo pelo qual duas formas podem ocorrer no mesmo contexto com o mesmo valor referencial/representacional, isto é, com o mesmo significado.” (COELHO *et al.*, 2018, p. 16). Por exemplo, na seção anterior, foi mostrada a alternância das formas *nós* e *a gente* na codificação da segunda pessoa do plural no PB. Às duas formas pronominais em variação dá-se o nome de variantes, enquanto à função gramatical em que se localiza a variação de forma mais abstrata, chama-se de variável.

A partir dos estudos labovianos (LABOV, 1972, 1978, 1994, 2001), descarta-se a ideia de a variação ocorre de forma aleatória, uma vez que o autor mostra que forças, externas ou internas, agem sobre a língua favorecendo ou desfavorecendo o uso de determinada variável. A essas forças podemos chamar de fatores de controle. Dentro desse contexto, as pesquisas em Sociolinguística Variacionista visavam ao estudo da estratificação sociolinguística da variação, ou seja, a investigação busca o encaixamento dos fenômenos linguísticos dentro da estrutura social. Porém, estudos como Weiner e Labov (1977) sobre a passiva sem agente mostraram que nem sempre os fatores extralinguísticos assumem papel relevante no condicionamento de uso de uma variável. A partir desse resultado, Lavandera (1978) questiona sobre o rótulo de “social” endereçado à Sociolinguística e sobre o lugar das variáveis sociais no estudo da

variação linguística. Em resposta a esse questionamento, Labov (1978) diz que os interesses das pesquisas sociolinguísticas se voltam igualmente às restrições internas do sistema e que as análises dos dados se constituem como mecanismos para saber sobre a gramática.

Com base nessa visão, este estudo foca na análise dos fatores linguísticos no condicionamento da alternância dos pronomes sujeitos *nós* e *a gente*. Para entender tal variação e o avanço da forma *a gente* na fala dos fortalezenses na segunda década dos anos 2000, é importante discutirmos a noção de norma, considerando, sobretudo, a diferença entre norma padrão e norma culta. Faraco (2008) explica que ambas são associadas ao conjunto de usos linguísticos atrelados ao grupo letrado da sociedade, porém a primeira se relaciona à ideia de normatividade, enquanto a segunda à ideia de normalidade.

Para o autor, a norma padrão é um modelo de língua relativamente abstrato, em que uma determinada variedade é usada como referência na tentativa de uma política de uniformização linguística. Já a norma culta, objeto deste estudo, diz respeito ao conjunto dos fenômenos linguísticos (variáveis), os quais ocorrem no uso dos falantes de nível superior em situações monitoradas de fala e de escrita.

Trazemos ainda, dentro dessa perspectiva, a conceituação de variantes padrão ou não-padrão; de prestígio ou estigmatizada e conservadoras ou inovadoras. Coelho *et al.* (2018) explica que as variantes padrão condizem com os padrões normativos da língua e as não padrão distanciam-se deles. Geralmente, as primeiras trazem consigo prestígio social, enquanto a segunda é estigmatizada. Ademais, as conservadoras são aquelas que estão há mais tempo no sistema linguístico da comunidade - como é o caso do pronome *nós* - diferente das inovadoras - como é o caso do pronome *a gente* - que são mais recentes na história da língua.

Foi com base nos conceitos supracitados que a metodologia deste estudo foi construída. Desse modo, a próxima seção apresenta o percurso metodológico da pesquisa.

Metodologia

Nossa pesquisa se insere no campo da Sociolinguística Quantitativa, portanto, trata-se de uma pesquisa de descrição linguística com tratamento estatístico de dados. Os dados analisados foram coletados de entrevistas orais que compõem o banco de dados do Projeto Descrição do Português Oral Culto de Fortaleza-CE (PORCUFORT) - Fase II (FREITAS; RODRIGUES; SANTOS, 2021). O banco de dados teve sua primeira fase idealizada pelo professor José Lemos Monteiro e atualmente encontra-se sediado na Universidade Estadual do Ceará sob a tutela da professora Aluiza Alves Araújo.

O PORCUFORT conta com inquéritos de três tipos de registro: Diálogo entre Informante e Documentador (DID), Diálogo entre Dois Informantes (D2) e Elocuções Formais (EF). Além

do tipo de registro, a estratificação dos informantes, fortalezenses com nível superior, é feita por sexo (masculino e feminino) e faixa etária (I - 22 a 35 anos; II - 36 a 55 anos; III - 56 anos ou mais). A amostra selecionada em nossa pesquisa conta com 24 inquiridos do PORCUFORT - Fase II da modalidade de registro DID, seguindo a estratificação por sexo e faixa etária, portanto, contamos com quatro informantes por célula social. Para este trabalho, no entanto, analisou-se apenas dados do tipo de registro DID, sendo selecionado três informantes por célula, conforme apresentado no Tabela 1.

Tabela 1: Distribuição dos informantes na amostra analisada

FAIXA ETÁRIA	GÊNERO	
	Masculino	Feminino
I	3	3
II	3	3
III	3	3
Total	9	9
	18	

Fonte: elaborado pelos autores a partir de Araújo, Viana e Pereira (2018).

Em relação às variáveis linguísticas⁴ controladas, analisou-se o tempo e tipo de paradigma do verbo, a referencialidade do pronome (genérica ou específica) e o tipo de verbo.

O tratamento estatístico dos dados foi feito com o *software* RStudio. Com essa ferramenta, procedemos à análise de estatística descritiva, com aplicação de testes de qui-quadrado às distribuições encontradas das variantes em estudo por variável controlada, e à análise de estatística inferencial, com a elaboração de modelos de regressão logística para análise dos dados linguísticos. Após as rodadas estatísticas, os dados foram analisados e, na seção a seguir, discutiremos esses resultados.

Análise dos resultados e discussão

Iniciando pela proporção geral de dados, em nossa amostra obtivemos um total de 1197 dados de *nós e a gente*. A tabela a seguir apresenta a distribuição dos nossos dados:

Tabela 2: Distribuição das variantes na amostra completa (N = 1197)

	Ocorrências	%
Nós	213	17,78
A gente	985	82,22

⁴ Na segunda seção, esses fatores foram apresentados detalhadamente.

Analisamos a alternância entre as duas variantes linguísticas e constatamos uma preferência pela forma *a gente* com um total de 82,22% dos dados, com um total de 985 ocorrências, sendo a forma *nós* preterida, representando 17,78% dos dados e 213 ocorrências. Outros estudos sociolinguísticos já haviam atestado dados acima de 80% da forma *a gente* desde a década de 1980 na fala de pessoas com ensino superior em dados do projeto NURC (ZILLES, 2007), contudo Fortaleza-CE em dados dos anos 90 tinha uma predominância de “apenas” 62,01% da forma *a gente* (CARVALHO; FREITAS; FAVACHO, 2020). Com isso, observamos um grande avanço da forma inovadora na fala de Fortaleza, tendo uma predominância massiva dos dados de primeira pessoa do plural.

A partir desse conjunto de dados, operamos um recorte, buscando sanar as células vazias observadas na distribuição de dados pela variável tempo e tipo de paradigma verbal. Dessa forma, foram selecionadas as ocorrências que envolviam apenas os tempos do modo indicativo. Esse recorte resultou na exclusão de 80 ocorrências para as rodadas que apresentamos a seguir. Portanto, a partir de agora, utilizaremos uma subamostra de 1117 ocorrências.

A tabela a seguir demonstra a distribuição das variantes em análise por tempo e tipo de paradigma verbal:

Tabela 3: Distribuição das variáveis pelo tempo verbal no modo indicativo (N = 1117)

Tempo verbal	Nós		A gente	
	Ocorrências	%	Ocorrências	%
Presente igual	11	5,24	199	94,76
Presente diferente	61	20,61	235	79,39
Pretérito igual	63	37,72	104	62,28
Imperfeito	31	9,90	282	90,10
Pretérito diferente	42	32,06	89	67,92
$X^2 = 46.457 (4), p < 0.001$				

A variável tempo e tipo de paradigma do verbo, como já dito, é uma das variáveis linguísticas mais importantes para descrição da variação *nós* e *a gente* devido ao seu alto poder explicativo. Nosso teste de qui-quadrado demonstrou um valor de 46,457 com quatro graus de liberdade, ou seja, devido a variável possuir cinco níveis, existem sempre 4 outras possibilidades de tempo verbal a cada variante. Em um teste de qui-quadrado, os valores da distribuição encontrada são comparados aos valores da distribuição esperada entre as variantes. Assim, quanto mais distante de zero, maior a probabilidade de a distribuição de

dados encontrada ser resultado da atuação de algum fator da variável. Em nosso caso, o valor de qui-quadrado encontrado é distante de zero e o valor $p < 0,001$ reforça a percepção de que o teste aponta para a influência dessa variável sobre a variação *nós e a gente*. Dessa forma, temos indícios de que essa variável será selecionada na regressão logística.

Vemos que, em termos de proporção e frequência, a forma *a gente* é predominante em todos os tempos verbais, um fator positivo no processo de mudança linguística, devido a possibilidade de uso em todos os níveis da variável. Percebemos uma maior proporção de *a gente* no pretérito imperfeito e no tempo presente, fatores tradicionalmente selecionados como favorecedores a forma inovadora (NARO; GÖRSKI; FERNANDES, 1999; SCHERRE; YACOVENCO; NARO, 2018a, 2018b). Entretanto, aqui tratamos apenas em termos descritivos de estatística básica e proximamente nesse texto discutiremos esses dados a partir da estatística inferencial dos modelos de regressão logística para aprofundamento da análise.

A próxima variável controlada em nossa pesquisa foi a referência do pronome. A tabela a seguir indica a distribuição encontrada:

Tabela 4: Distribuição das variáveis por referência (N = 1117)

Referência	Nós		A gente	
	Ocorrências	%	Ocorrências	%
Genérica	42	9,29	410	90,71
Específica	166	24,96	499	75,06

$X^2 = 7.5819 (1), p = 0.005896$

Pelas proporções encontradas, nossos dados indicam um favorecimento da forma inovadora em contextos de referência genérica (90,71%) enquanto a forma tradicional é favorecida pela referência específica (24,96%). Os dados atestam uma predominância da forma *a gente* tanto na referência mais genérica quanto a mais específica, ainda assim podemos perceber uma maior predominância da forma *a gente* como mais genérica e uma presença mais marcante do pronome *nós* como forma específica, ou seja, demonstra que o matiz semântico influencia o uso de uma variante em detrimento de outra. O teste de qui-quadrado aplicado a essa distribuição resultou em um valor de 7,5819, com um grau de liberdade, e um p-valor $< 0,05$. Esses resultados indicam que essa variável pode ser selecionada em uma regressão logística.

A terceira variável linguística controlada em nossa pesquisa foi o tipo de verbo. A tabela a seguir indica a distribuição das variantes para essa variável:

Tabela 5: Distribuição das variáveis por tipo de verbo (N = 1117)

Tipo de verbo	Nós		A gente	
	Ocorrências	%	Ocorrências	%
Dicendi	2	6.25	30	93.75
Ter	35	20.35	137	79.65
Ação	128	17.20	616	82.80
Epistêmico	7	13.73	44	86.27
Estado	36	30.51	82	69.49

$X^2 = 21.933 (4), p < 0.001$

De acordo com a distribuição encontrada, a variante *a gente* seria favorecida por verbos dicendi (93,75%), epistêmicos (86,27%) e de ação (82,8%). Já a forma *nós* seria favorecida pelos verbos de estado (30,51%), pelo verbo ter (20,35%) e por verbos de ação (17,2%). É interessante observar que a ordenação dos fatores para cada variante é inversa, ou seja, poderíamos traçar um *continuum* entre os tipos de verbos. Para a forma *a gente*, observamos a seguinte sequência de favorecimento pelo tipo de verbo: Dicendi > Epistêmico > Ação > Ter > Estado. Para a forma *nós*, a sequência é a seguinte: Estado > Ter > Ação > Epistêmico > Dicendi.

O teste de qui-quadrado dessa variável apresenta um valor de 21,933, com quatro graus de liberdade e p-valor < 0,01. Portanto, há indícios de que essa variável exerça influência sobre o fenômeno variável em estudo.

A partir desses dados, partimos para a elaboração de um modelo de regressão logística para explicitar os mecanismos que atuam na variação entre os pronomes *nós* e *a gente*. A tabela a seguir indica o resultado obtido com o teste do modelo que incluiu as três variáveis: (1) referência, (2) tempo verbal e tipo de paradigma e (3) tipo de verbo. Os valores desse modelo se referem às probabilidades de uso da forma inovadora *a gente*. Vejamos a tabela:

Tabela 6: Resultados do modelo de regressão logística (N = 1117)

Coeficientes	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	3.71898	0.88607	4.197	4.197	
Referência					
Específica	-0.95380	0.25486	-3.742	0.000182	***
Tempo verbal					
Presente dif.	-1.27102	0.37873	-3.356	0.000791	***
Pretérito igual	-1.72369	0.37727	-4.569	4.90e-06	***
Imperfeito	0.20587	0.39699	0.519	0.604051	
Pretérito dif.	-1.39641	0.39255	-3.557	0.000375	***
Tipo de verbo					
Ter	-1.23886	0.82341	-1.505	0.132442	
Ação	-0.42252	0.79315	-0.533	0.594239	
Epistêmico	-0.83888	0.90888	-0.923	0.356017	
Estado	1.51077	0.82031	1.842	0.065516	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					
glm = VD ~ referencia + tempoverbal + tipodeverbo					

Inicialmente, é necessário compreender de que tratam o *intercept* e a estimativa. O valor de *intercept* se refere à probabilidade de emprego da variante *a gente* no primeiro nível das variáveis linguísticas controladas. Assim, a estimativa de *intercept* corresponde ao uso de *a gente* com (1) referência genérica, (2) no presente do indicativo com verbo de paradigma igual e (3) do tipo dicendi.

Para averiguar a atuação de cada um dos outros fatores das variáveis, é preciso somar o valor das estimativas correspondentes. Por exemplo, para ver a atuação do fator específico da variável referência, somamos a estimativa de *intercept* (3,71898) à estimativa do fator específico (-0,95380). O resultado dessa soma (2,76518) indica que a probabilidade de emprego da forma *a gente* com referência específica, mantidos os outros fatores, é menor do que a probabilidade indicada inicialmente, com referência genérica. Em outras palavras, o modelo indica que a referência específica desfavorece o uso da forma *a gente*.

A coluna das estimativas, como é possível inferir da explanação sobre o *intercept*, é uma medida estatística da probabilidade em *logodds* do emprego da variante focalizada na análise de

regressão logística. O sinal positivo ou negativo que acompanha o valor de estimativa de cada fator pode ser tomado como indicativo de que aquele fator atua favorecendo ou não o emprego da variante.

A coluna do desvio padrão e do valor-z são medidas estatísticas que indicam a variabilidade dos dados. Essas medidas não são consideradas em nosso estudo. A última coluna, no entanto, apresenta o p-valor associado a cada fator de cada variável seguido de asteriscos que situam esses valores numa escala de 0 a 1. Quanto mais próximo de zero, maiores são os indícios de atuação do fator na variação.

Feitas essas considerações, vamos às conclusões que podemos retirar desse modelo. Das três variáveis consideradas no modelo, apenas a variável tipo de verbo não foi selecionada como estatisticamente relevante. Embora o teste de qui-quadrado tenha apontado indícios de que essa variável poderia atuar na variação entre os pronomes *nós* e *a gente*, o modelo de regressão logística testa a atuação das variáveis conjuntamente. Dessa forma, o tipo de verbo não se mostrou uma variável forte para esse fenômeno.

Sobre a atuação da variável referência, como já destacado, foi observada uma tendência de emprego da variante inovadora em contexto de referência genérica. Acreditamos que o principal motivo seja o traço genérico e menos determinado originário de sua forma nominal anterior ao processo de gramaticalização que se manteve como matiz semântico (LOPES, 2007), ainda que o traço mais específico também haja sido assimilado pela variante *a gente* esse condicionamento linguístico pode ser explicado pela história da variante. Outro fator que acreditamos que pode ser imbricado é, como comentado, a referência original de terceira pessoa do singular, acreditamos que existe a possibilidade de que a forma *nós* esteja se especializando como designadora de especificidade, considerando a reorganização do paradigma pronominal do português (FARACO, 2017) e a simplificação de desinências verbais, sendo a desinência zero, original da terceira pessoa do singular (*ele vai*), ampliando-se para também segunda pessoa do singular (*você vai*) e primeira pessoa do plural (*a gente vai*), dessa maneira, utilizar da forma de primeira pessoa do plural *nós* com *-mos* pode ser um artifício para enfatizar e especificar a referência do pronome.

Sobre a variável tempo e tipo de paradigma verbal, o único fator não selecionado foi o pretérito imperfeito. O tempo pretérito imperfeito tem sido selecionado como favorecedor da forma *a gente* sem *-mos* em diversas pesquisas (LOPES, 1998, 2007; SCHERRE; YACOVENCO; NARO, 2018a, 2018b; CARVALHO; FREITAS; FAVACHO, 2020, para citar apenas alguns estudos). Apesar do modelo de regressão não haver selecionado este nível da variável, devemos considerar que o total de 282 ocorrências de *a gente* com pretérito perfeito, totalizando 90,10% de preferência dos dados. Portanto, não podemos ignorar que esse é um fator tradicionalmente favorecedor de *a gente*, em uma possível esquia as proparoxítonas tradicionais da desinência canônica de primeira pessoa do plural (*estávamos/comíamos*), e

percebemos esse favorecimento no nosso *corpus* com uma alta preferência ao uso de *a gente*, ainda que não tenha sido selecionado pelo modelo.

Os outros fatores apresentam a seguinte ordenação: presente igual > presente diferente > pretérito diferente > pretérito igual. A hipótese plantada em Naro, Görski, Fernandes (1999) é de que a desinência *-mos*, associada ao uso de *nós* devido aos altos índices de concordância verbal normalmente apresentados nas amostras do Português Brasileiro, seria um recurso para enfatizar o pretérito, e a forma *a gente* sem *-mos* seria utilizada para marcar o presente. Ainda que essas sejam desinências de pessoa verbal, não semântico-temporais, o uso de uma forma em detrimento de outra parece ter se especializado para evitar a ambiguidade do sistema desinencial. Com isso, observamos em nossos dados o contexto quase categórico de *a gente* com forma igual ao pretérito perfeito (94,76%), e o maior desfavorecimento da forma *a gente* para o pretérito perfeito igual ao presente na regressão logística ($p= 4.90e-06$ ***). Devemos ter em conta que o presente, mesmo sendo historicamente favorecido pelo *a gente*, quando diferente do presente, ainda que proporcionalmente tenha sido de maior ocorrência da forma inovadora (79,39%), o modelo de regressão logística apresentou uma estimativa negativa para o uso dessa forma (estimate = -1.27102), o que significa desfavorecimento. Esse resultado pode levar ao raciocínio que, em nossos dados, o maior condicionador de uma forma em detrimento de outra foi a ambiguidade temporal presente/pretérito perfeito.

Por fim, cabe ainda destacar o índice de Concordância C do modelo de regressão elaborado. Essa medida estatística diz respeito ao poder explicativo do modelo. O índice C de nosso modelo é 0,821, valor considerado por Levshina (2015) como indicativo de notório poder explicativo do modelo.

Considerações finais

Esta pesquisa teve como objetivo analisar o efeito de fatores linguísticos sob a alternância do uso das variantes de primeira pessoa do plural *nós* e *a gente* na fala de falantes de Fortaleza com ensino superior na segunda década dos anos 2000, a partir de dados do projeto Português Oral Culto de Fortaleza (PORCUFORT) em sua fase II. Consideramos os pressupostos teórico-metodológicos da Teoria de Variação e Mudança Linguística (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 2008 [1972]) para investigar a possibilidade de mudança em curso, compreender o encaixamento das variantes no sistema linguísticos e possíveis condicionadores ao uso de uma forma em detrimento de outra.

As variáveis estudadas aqui foram: tempo e tipo de paradigma do verbo no modo indicativo, referência do verbo (genérica ou específica) e tipo de verbo. Para o tempo e tipo de paradigma do verbo, o modelo de regressão logística indicou a predominância da forma *a gente* no presente de forma idêntica a pretérito perfeito na desinência de primeira pessoa do plural (comemos *ontem*/ comemos *hoje*). A variante também foi mais desfavorecida no pretérito

perfeito de forma idêntica ao presente, o que indica o favorecimento da forma conservadora *nós* nesse contexto linguístico. Tal resultado está em consonância ao estudo de Naro, Görski, Fernandes (1999) de que o uso das variantes poderia ser especializado para ajudar a desfazer a potencial ambiguidade semântico-temporal intrínseca ao paradigma verbal da desinência de primeira pessoa do plural.

Para referência do verbo, considerarmos dois níveis, a referência genérica, traço semântico [+ indeterminado] e a específica, traço semântico [+ determinado]. As frequências e proporções, em termos de estatística descritiva, apontam uma predominância de *a gente* em ambos os contextos, contudo o modelo de regressão logística aponta que a forma inovadora é desfavorecida pela referência específica, fato possivelmente relacionado a sua origem nominal e processo de gramaticalização, em que seria mantido um matiz semântico de indeterminação da forma original (LOPES, 2007). A forma *nós* poderia estar especializando-se para demarcar a referência mais específica, aproveitando-se também da desinência verbal de primeira pessoa do plural, perdida no caso da forma *a gente*, que concorda com a desinência zero, originalmente de terceira pessoa do singular. Ou seja, utilizar a forma *nós* com concordância poderia ser um recurso para enfatizar a determinação do sujeito.

O tipo de verbo obteve, como no tópico anterior, maior predominância da variante *a gente* em todos os níveis (estado-epistêmico-*dicendi-ter-ação*) não se demonstrou relevante no modelo de regressão logística, ainda que tenha sido observado um desfavorecimento mais notório da forma *a gente* em verbos de estado.

Acreditamos que existe um processo de mudança em curso em direção a forma inovadora *a gente* devido a sua alta implementação no sistema linguístico, com 82,22% dos dados, e uma predominância, em termos de frequência, em todos os níveis das variáveis linguísticas analisadas, sendo os condicionadores positivos ou negativos ao seu uso apontados nos modelos de regressão logística. Com isso, observamos o avanço da forma e seu uso autorizado e predominante em todos os contextos analisados e um cenário que a forma conservadora parece resistir apenas em contextos especializados de marcação de pretérito perfeito e referência específica.

Referências

ARAÚJO, Aluiza Alves de; VIANA, Rakel Beserra de Macêdo; PEREIRA, Maria Lidiane de Sousa. O projeto descrição do português oral culto de Fortaleza - PORCUFORT: das origens aos dias atuais. *Web-Revista SOCIODIALETO* - NUPESDD/LALIMU, v. 8, n. 24, mar. 2018. Disponível em: <http://sociodialeto.com.br/index.php/sociodialeto/article/view/39/91>. Acesso em: 03 maio 2021.

ARAÚJO, Marden Alyson Matos de. Nós e a gente no falar dos fortalezenses. In: ARAÚJO, Aluiza Alves de; VIANA, Rakel Beserra de Macêdo; PEREIRA, Maria Lidiane de Sousa. (org.). *Fotografias sociolinguísticas do falar de Fortaleza – CE*. Fortaleza: EdUECE, 2018. p. 143-172.

CARVALHO, Hebe Macedo de; FREITAS, Maylle Lima; FAVACHO, Larissa de Lima. A variação dos pronomes sujeitos nós e a gente: a fala culta de Fortaleza em cena. *Revista (Con) textos linguísticos*, v. 14, n. 27, Espírito Santo, 2020. Disponível em: <https://periodicos.ufes.br/contextoslinguisticos/article/view/29213>. Acesso em: 25 mar. 2022.

COELHO, Izete Lehmkuhl *et al.* *Para conhecer sociolinguística*. São Paulo: Editora Contexto, 2018.

FARACO, Carlos Alberto. *Norma Culta Brasileira: desatando alguns nós*. São Paulo: Parábola Editorial, 2008.

FARACO, Carlos Alberto. O tratamento “você” em português: uma abordagem histórica. *LaborHistórico*, Rio de Janeiro, vol.3, n. 2. Rio de Janeiro, p. 114-132, 2017. Disponível em: <https://revistas.ufrj.br/index.php/lh/article/view/17150>. Acesso em: 10 nov. 2021.

FREITAG, Raquel Meister Ko. Uso, crença e atitudes na variação na primeira pessoa do plural no português brasileiro. *DELTA*, v. 32, n. 4, p. 889-917, 2016. Disponível em: <https://revistas.pucsp.br/index.php/delta/article/view/29225/22229>. Acesso em: 21 set. 2021.

FREITAS, Maylle Lima; CARVALHO, Hebe Macedo de. Quem somos “nós” e quem é “a gente”? Uma abordagem de avaliação linguística e social da variável de primeira pessoa plural. In: VIANA ; RAKEL Beserra de Macêdo; RODRIGUES, Lorena da Silva; PONTES, Valdecy de Oliveira; CARVALHO, Hebe Macedo de (org.). *Estudos em sociolinguística variacionista e sociofuncionalismo*. São Paulo: Pimenta Cultural, 2020, p. 124-142. DOI: 10.31560/pimentacultural/2020.985.124-142.

FREITAS, Maylle Lima; FAVACHO, Larissa de Lima; CARVALHO, Hebe Macedo de. Avaliação e percepção das formas *nós* e *a gente* e dos padrões de concordância por falantes escolarizados de Fortaleza -Ceará. *Miguilim–Revista Eletrônica do Netlli*, Crato, v. 11, n. 1, p. 40-57, jan.-abr. 2022.

FREITAS, Maylle Lima; RODRIGUES, Lorena da Silva; SANTOS, Hugo Leonardo Gomes dos. *Nós e a gente* no falar culto de Fortaleza: variação ou mudança linguística?. In: ARAÚJO, Aluiza Alves de; RODRIGUES, Lorena da Silva; VIANA, Rakel Beserra de Macêdo. *O falar culto de Fortaleza em foco*. São Paulo: Pimenta Cultural, 2021. p. 139-171. DOI: 10.31560/pimentacultural/2021.834.139-171.

LABOV, William. *Padrões sociolinguísticos*. São Paulo: Parábola, 2008.

LABOV, William. Where does the linguistic variable stop? A response to Beatriz Lavandera. In: *Sociolinguistics working paper*. Austin: Southwest Educational Development Laboratory, n. 44, 1978.

LABOV, William. *Principles of linguistic change: internal factors*. Cambridge, MA: Blackwell, 1994.

LABOV, William. *Principles of linguistic change: social factors*. Oxford: Blackwell, 2001.

Diadorim, Rio de Janeiro, vol. 24, número 1, p. 161 - 179, 2022.

LAVANDERA, Beatriz. Where does the linguistic variable stop? *Language society*. 7. London, 1978.

LEVSHINA, Natalia. *How to do linguistics with R: data exploration and statistical analysis*. Philadelphia: John Benjamins Publishing Co., 2015.

LOPES, Célia Regina dos Santos. Nós e a gente no português falado culto do Brasil. *DELTA*, v. 14, n. 2, p. 1-12, 1998. Disponível em: <https://revistas.pucsp.br/index.php/delta/article/view/44300>. Acesso em: 22 set. 2021.

LOPES, Célia Regina dos Santos. A gramaticalização de *A gente* em português em tempo real de longa e de curta duração: retenção e mudança na manutenção dos traços intrínsecos. *Fórum Lingüístico*, Florianópolis, v. 4, n. 1, p. 47-80, 2007. Disponível em: <https://periodicos.ufsc.br/index.php/forum/article/view/7728>. Acesso em: 10 nov. 2021.

LUCCHESI, Dante. A representação da primeira pessoa do plural. In: LUCCHESI, Dante; BAXTER, Alan; RIBEIRO, Ilza (org.). *O português afro-brasileiro*. Salvador: Ed. da UFBA, 2009. p. 457-468.

MAIA, Francisca Paula Soares. A variação Nós/ A gente no dialeto mineiro: investigando a transição. *Revista da ABRALIN*, v. 8, n. 2, p. 45-70, 2009. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/1032>. Acesso em: 11 nov. 2021.

NARO, Anthony; GÖRSKI, Edair; FERNANDES, Eulália. Change without change. *Language Variation and Change*, v. 11, p. 197-211, 1999.

OMENA, Nelize Pires de. A referência à primeira pessoa do plural. In: SILVA, Giselle Machline de Oliveira e; SCHERRE, Maria Marta Pereira (orgs.). *Padrões sociolinguísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro, 1996. p. 183-216.

OMENA, Nelize Pires de. A referência à primeira pessoa do plural: variação ou mudança? In: PAIVA, Maria da Conceição; DUARTE, Maria Eugênia Lamoglia (orgs.). *Mudança linguística em tempo real*. Rio de Janeiro: Contra Capa, 2003. p. 63-80.

OUSHIRO, Lívía. Tratamento de Dados com o R para Análises Sociolinguísticas. In: FREITAG, Raquel Meister ko (org.). *Metodologia de coleta e manipulação de dados em Sociolinguística*. São Paulo: Edgard Blücher Ltda, 2014. p. 129-172.

R CORE TEAM. *The R Project for Statistical Computing*. Página Web. Disponível em: <https://www.r-project.org/>. Acesso em: 20 jun. 2021.

SILVA, Francisca Jocineide de Alencar. *A variação nós e a gente na fala de Fortaleza*. 2020. 91f. Dissertação (mestrado em linguística) - Programa de Pós-Graduação em Linguística (PPGLin). Universidade Federal do Ceará, Fortaleza, 2020.

SCHERRE, Maria Marta Pereira; YACOVENCO, Lilian Coutinho; NARO, Anthony Julius. Nós e a gente no português brasileiro: concordâncias e discordâncias. *Estudos de Lingüística Galega*, v. especial I, ed. F. Cidrás, F. Dubert and X. L. Regueira, p. 13-27, 2018a. Disponível em: <https://www.cambridge.org/core/journals/language-variation-and-change/article/abs/socio-linguistic-correlates-of-negative-evaluation-variable-concord-in-rio-de-janeiro/B209D52E57FC42496C947A105A6A3B55>. Acesso em: 22 fev. 2021.

SCHERRE, Maria Marta Pereira; YACOVENCO, Lilian Coutinho; NARO, Anthony Julius. Nós e a gente em quatro amostras do português brasileiro: revisitando a escala da saliência fônica. *Diadorim*, v. 20 – Especial, p. 428-457, 2018b. Disponível em: <https://revistas.usc.gal/index.php/elg/article/view/3585>. Acesso em: 22 set. 2021.

VIANNA, Juliana Segadas; LOPES, Célia Regina dos Santos. Variação dos pronomes nós e a gente. In: MARTINS, Marco Antonio; ABRAÇADO, Jussara (org.). *Mapeamento sociolinguístico do português brasileiro*. São Paulo: Contexto, 2015. p. 109-131.

WEINER, E. Judith; LABOV, William. Constraints on the agentless passive. *Journal of Linguistics*, n.º. 19, 1977.

WEINREICH, Uriel; LABOV, William; HERZOG, Marvin. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola, 2006.

ZILLES, Ana Maria. O que a fala e a escrita nos dizem sobre a avaliação social do uso de “a gente”. *Letras de Hoje*, v. 42, n. 2, Porto Alegre, p. 27-44, 2007. Disponível em: <http://repositorio.furg.br/bitstream/handle/1/9133/DISSERTA%C3%87%C3%83O%20DE%20MESTRADO%20MATHEUS.pdf?sequence=1>. Acesso em: 21 set. 2021.



**O EFEITO DO PARALELISMO LINGUÍSTICO SOBRE A DISPUTA
ENTRE O INDICATIVO E O SUBJUNTIVO NA FORMAÇÃO DO
IMPERATIVO DE 2ª PESSOA DO SINGULAR NO PORTUGUÊS
MINEIRO HISTÓRICO (SÉCULOS XIX E XX)**

**THE EFFECT OF LINGUISTIC PARALLELISM ON THE DISPUTE
BETWEEN THE INDICATIVE AND THE SUBJUNCTIVE IN THE
FORMATION OF THE 2ND PERSON SINGULAR IMPERATIVE IN
HISTORICAL PORTUGUESE FROM MINAS GERAIS
(19TH AND 20TH CENTURIES)**

Luiz Fernando de Carvalho¹

RESUMO

No português brasileiro, o imperativo de 2ª pessoa do singular expressa-se de modo variável por meio de formas do indicativo (*deixa/recebe/abre*) e do subjuntivo (*deixe/receba/abra*) (FARACO, 1982; PAREDES SILVA *et al.*, 2000; SCHERRE, 2007; RUMEU, 2016; CARVALHO, 2020). Neste artigo, a partir de cartas mineiras oitocentistas e novecentistas autógrafas, mapeia-se um conjunto de dados de formas imperativas, com o objetivo de explorar o papel do paralelismo linguístico (SCHIFFRIN, 1981; WEINER & LABOV, 1983; SCHERRE, 1988; 1998) sobre a expressão variável. Desse modo, à luz da Sociolinguística Histórica (ROMAINE, 1982; HERNÁNDEZ-CAMPOY & CONDE SILVESTRE, 2012), inspirada, por seu turno, na Sociolinguística Variacionista (LABOV, 1972), atesta-se que a variação do imperativo nas missivas pesquisadas é sensível ao paralelismo sintático e ao paralelismo fônico, tendo em vista a seleção desses grupos de fatores como estatisticamente relevantes pelo programa GoldVarbX (SANKOFF, TAGLIAMONTE & SMITH, 2005). Esses resultados permitem compreender a atuação de condicionamentos linguísticos na dinâmica da variação do imperativo em amostras históricas.

PALAVRAS-CHAVE: Modo Imperativo; Paralelismo linguístico; Sociolinguística Histórica; Cartas pessoais.

¹ Doutorando do Programa de Pós-graduação em Estudos Linguísticos da Universidade Federal de Minas Gerais, lufecarva@gmail.com.

ABSTRACT

In Brazilian Portuguese, the 2ND person singular of the imperative mood is variably expressed through the forms of the indicative (*deixa/recebe/abre*) and the subjunctive (*deixe/receba/abra*) (FARACO, 1982; PAREDES SILVA et al., 2000; SCHERRE, 2007; RUMEU, 2016; CARVALHO, 2020). In this paper, on the basis of autographed eighteenth and nineteenth century letters from Minas Gerais, a dataset of imperative forms is mapped with the objective of exploring the role of linguistic parallelism (SCHIFFRIN, 1981; WEINER & LABOV, 1983; SCHERRE, 1988; 1998) on variable status. Thus, in the light of Historical Sociolinguistics (ROMAINE, 1982; HERNÁNDEZ-CAMPOY & CONDE SILVESTRE, 2012), inspired, in turn, by Variationist Sociolinguistics (LABOV, 1972), it is confirmed that the variation of the imperative in the researched missives is sensitive to syntactic and phonic parallelism, considering the selection of these groups of factors as statistically relevant by the GoldVarbX program (SANKOFF, TAGLIAMONTE & SMITH, 2005). These results allow us to understand the role of linguistic conditioning in the dynamics of imperative variation in historical samples.

KEYWORDS: Imperative Mood; Linguistic Parallelism; Historical Sociolinguistics; Personal Letters.

Considerações iniciais

Na 2ª pessoa do singular (doravante 2SG), o modo imperativo se expressa por formas verdadeiras, do indicativo (*deixa/recebe/abre/dá/diz/vai*), e supletivas, do subjuntivo (*deixe/receba/abra/dê/diga/vá*), promovendo o que se convencionou denominar, à luz de Scherre (2007, p. 190), imperativo verdadeiro e imperativo supletivo, nessa ordem. Na visão da tradição gramatical (BECHARA, 2009 [1961]; ROCHA LIMA, 2013 [1972]; CUNHA; CINTRA, 2007 [1985]), enquanto o imperativo verdadeiro é prescrito para contextos de *tu* em posição de sujeito, o imperativo supletivo tem seu emprego limitado aos contextos de sujeito *você*.

Todavia, no português brasileiro (doravante PB), as formas indicativa e subjuntiva apresentam um potencial variável condicionado por fatores linguísticos e extralinguísticos (FARACO, 1982; PAREDES SILVA et al., 2000; SCHERRE, 2003, 2007, 2012; CARDOSO, 2009; DINIZ, 2018; RUMEU, 2019; CARVALHO, 2020; entre outros). A variação desse modo verbal no PB é compreendida como uma das consequências da inserção do *você* que avançou sobre o espaço funcional do *tu* no sistema pronominal da língua (LOPES, 2007; LOPES & CAVALCANTE, 2011)², afetando a distribuição complementar entre as formas imperativas e sua referência ao sujeito de 2SG. Esse processo possibilitou uma manifestação *sui generis* do imperativo de 2SG no PB: o *imperativo abrasileirado*, entendido, nos termos de Paredes Silva et al. (2000, p. 121), como a combinação da forma indicativa em contexto de *você-sujeito*, como se exemplifica através do *slogan* comercial “*Vem pra Caixa Você também! Vem!*”³.

2 A tradição normativa considera que a posição de 2SG é ocupada tão somente pela forma pronominal *tu*. Todavia estudos linguísticos no âmbito da sociolinguística atestam que, na referida posição, há uma alternância entre as formas *tu* e *você* a depender de fatores linguísticos e extralinguísticos (LOPES, 2007, p. 116).

3 Slogan de uma conhecida campanha publicitária da Caixa Econômica Federal que pode ser verificada no endereço <https://brasilecola.uol.com.br/gramatica/intencionalidade-linguagem-publicitaria.htm>.

Neste artigo⁴, com base nos pressupostos teórico-metodológicos da Sociolinguística Histórica (ROMAINE, 1982; CONDE SILVESTRE, 2007; HERNÁNDEZ-CAMPOY & CONDE SILVESTRE, 2012), inspirada, por seu turno, na Sociolinguística Variacionista (WEINREICH; LABOV & HERZOG, 2006 [1968]; LABOV, 2008 [1972]), discute-se o estatuto de variação do imperativo de 2SG no PB a partir de cartas autógrafas de mineiros majoritariamente ilustres. As missivas, gênero marcado por relações distensas e íntimas, foram escolhidas em razão da possibilidade de se acessar vernáculo de tempos pretéritos (AGUILAR, 1998) através de manuscritos preservados graças à importância dada aos seus escreventes no cenário social e histórico de Minas Gerais. Nesse sentido, essa investigação busca depreender o papel do paralelismo linguístico (SCHIFFRIN, 1981; WEINER & LABOV, 1983; SCHERRE, 1988, 1998), nesse caso o sintático e o fonético, na expressão desse fenômeno em sincronias do passado (séculos XIX e XX). Com esses objetivos no horizonte, esta pesquisa parte das seguintes questões norteadoras:

(a) Considerando os estudos relacionados ao fenômeno variável em uma perspectiva histórica (SILVA, 2017; DINIZ, 2018), as formas imperativas presentes no *corpus* seriam predominantemente alinhadas indicativo (imperativo verdadeiro) ou ao subjuntivo (imperativo supletivo)?

(b) Considerando que o fenômeno é motivado por fatores linguísticos e extralinguísticos (SCHERRE, 2007; RUMEU, 2019; CARVALHO, 2020), em que medida o paralelismo sintático e o paralelismo fônico (SCHERRE, 1998; 2007; CARDOSO, 2009; DINIZ, 2018) constituiriam variáveis significativas na expressão do imperativo de 2SG nas cartas mineiras?

As hipóteses para essas questões são duas: (a) tendo em vista a inserção gradual do *você* no sistema pronominal (LOPES, 2007; LOPES & CAVALCANTE, 2011), conjectura-se que os mineiros se expressem preferencialmente pelas formas do imperativo supletivo às formas do imperativo verdadeiro, como um reflexo da atuação da norma-padrão (BECHARA, 2009 [1961]; ROCHA LIMA, 2013 [1972]; CUNHA & CINTRA, 2007 [1985]) sobre os escreventes; (b) tendo como base estudos sobre o efeito do paralelismo sobre a variação do imperativo (SCHERRE, 1998; 2007; CARDOSO, 2009; DINIZ, 2018), presume-se que o paralelismo sintático e o fonético sejam fatores selecionados como estatisticamente significativos para explicar o fenômeno.

A fim de verificar essas hipóteses, são analisadas, à luz da Sociolinguística Histórica (ROMAINE, 1982; CONDE SILVESTRE, 2007; HERNÁNDEZ-CAMPOY & CONDE SILVESTRE, 2012), 202 cartas mineiras oitocentistas e novecentistas com dados do imperativo de 2SG distribuídas em acervos públicos. As ocorrências identificadas são devidamente computadas pelo GoldVarbX (SANKOFF; TAGLIAMONTE & SMITH, 2005), um confiável

4 O artigo constitui um recorte com alguns resultados obtidos por Carvalho (2020) em sua pesquisa para a obtenção do título de mestre pela Universidade Federal de Minas Gerais (UFMG).

programa para cálculo estatístico dos fatores responsáveis pela aplicação de uma regra variável, comumente utilizado em pesquisas de natureza sociolinguística.

Assim, com o intuito de depreender a expressão variável do imperativo de 2SG em função do paralelismo linguístico, este artigo se divide em 6 seções. A partir das considerações iniciais, com apresentação do fenômeno variável bem como das questões e hipóteses, é realizada, na primeira seção, uma descrição dos critérios semânticos, morfológicos e sintáticos do imperativo adotados neste estudo. Na segunda, são trazidas à cena algumas considerações sobre a variável paralelismo linguístico seguidas de alguns resultados dessa variável em pesquisas sociolinguísticas a respeito do tema. Em seguida, na terceira seção, além da abordagem dos procedimentos teórico-metodológicos da Sociolinguística Histórica, apresenta-se, sucintamente, o processo de quantificação computacional utilizado para obtenção dos resultados. Na sequência, na quarta, na quinta e na sexta seção, procede-se à análise dos resultados gerais bem como em relação ao paralelismo sintático e o paralelismo fônico. Por fim, nas considerações finais, procura-se responder às questões propostas e tecer algumas generalizações sobre o fenômeno variável.

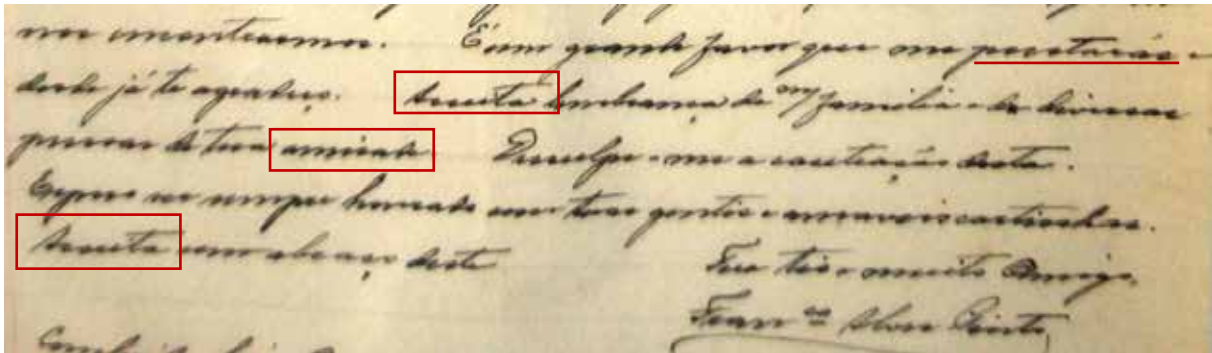
Os aspectos semânticos, morfológicos e sintáticos do imperativo de 2ª pessoa do singular

O imperativo, marcado por um ato ilocucionário (SEARLE, 1969 *apud* FARIA 2006, p. 73-74), manifesta-se em situações em que o interlocutor é levado a performar uma ação (pedido, ordem, súplica) proposta pelo locutor. Essa força ilocucionária está salvaguardada pelos aspectos formais que envolvem a expressão desse modo verbal. Assim, a fim de descrever o imperativo de 2SG, abordam-se, conforme alguns estudos norteadores (SCHERRE, 2007; CARDOSO, 2009; RUMEU & CARVALHO, 2018), os critérios morfossintáticos adotados neste artigo. Para tanto, serão utilizados os excertos (01), (02) e (03)⁵, transcritos das mãos dos missivistas mineiros, acompanhados de seus respectivos fac-símiles nas imagens (01), (02) e (03).

(01) É um grande favor que me prestarás e desde já te agradeço. *Acceita* lembrança de minha família e de diversas pessoas de tua amisade. *Desculpa*-me a [...] desta. Espero ser sempre honrado com tuas gentis e amaveis cartinhas. *Aceita* um abraço deste Teu tio e muito Amigo, Francisco Alves Pinto (FAPJ. Caeté, 19.08.1917) – *imperativo verdadeiro*

5 Nos exemplos deste estudo transcritos das cartas examinadas, optou-se por demarcar as formas imperativas em itálico e a referência ao contexto da variação com um grifo. Também, optou-se por preservar a identidade dos missivistas, exibindo sua autoria tão somente pelas iniciais de seus nomes.

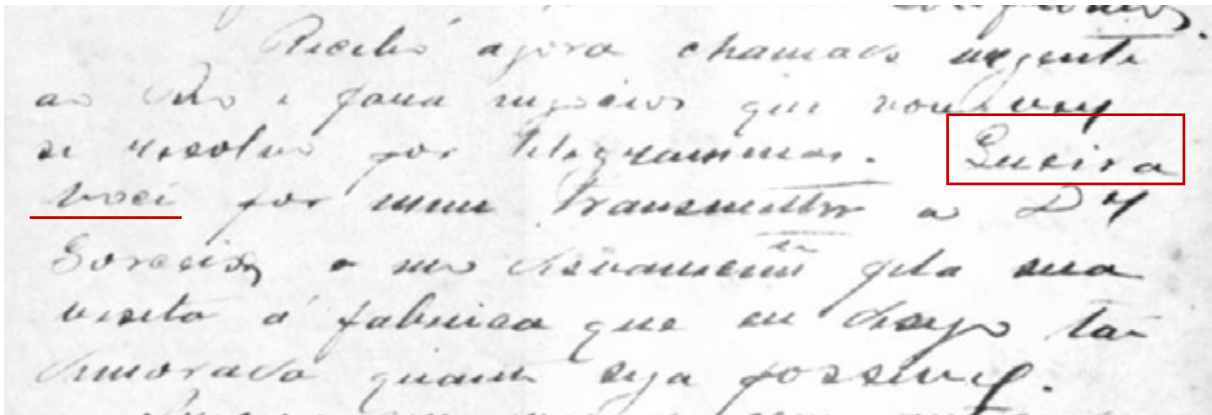
Imagem 01 – Carta de FAPJ. Caeté, 19.08.1917



Fonte: Instituto Histórico e Geográfico de Minas Gerais (IHGMG)

(02) Recebo agora chamado urgente ao Rio e para negocios que vou ver se resolvo por telegrammas. *Queira* você por mim *transmittir* ao Doutor Gorceix o meu desvanecimento pela sua visita á fabrica que eu desejo tão demorada quanto seja possível. (JP. Caeté, 30.12.1904) – *imperativo supletivo*

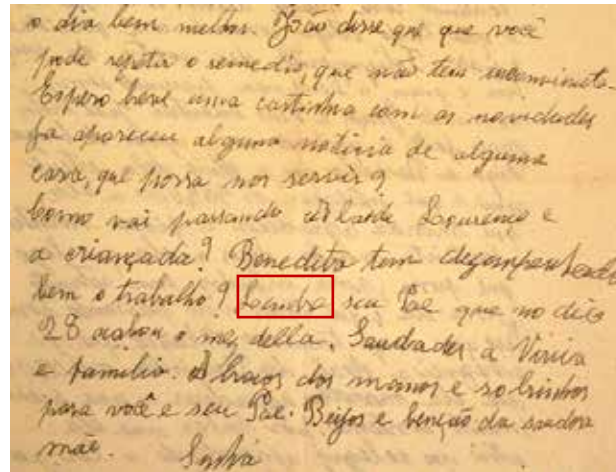
Imagem 02 – Carta de JP. Caeté, 30.12.1904



Fonte: Arquivo Público Mineiro (APM)

(03) João disse que você pode repetir o remédio, que não tem inconveniente. Espero breve uma cartinha com as novidades. Já apareceu alguma notícia de alguma casa, que possa nos servir? Como vai passando Alaide Lourenço e a criançada? Benedito tem desempenhado bem o trabalho? *Lembra* seu Pae que no dia 28 acabou o mez della. Saudades a Uzica e familia. Abraços dos manos e sobrinhos para você e seu Pae. (MRVL. s/ local, 02.02.1946) – *imperativo abraçador*

Imagem 03 – Carta de MRVL. s/ local, 02.02.1946



Fonte: Acervo dos Escritores Mineiros (AEM)

Enquanto em (01), as formas *aceita*, que se repete, e *desculpa* manifestam um pedido de um locutor direcionado ao interlocutor na despedida de uma carta⁶, em (02), o missivista, por meio da expressão *queira transmitir* requer que seu destinatário repasse ao Doutor Gorceix seu orgulho pela visita à fábrica. Por fim, em (03), a partir da forma *lembra*, a escrevente pede que seu correspondente recorde o pai do término do mês de Alaíde. Assim, na medida em que nessas ocorrências o destinatário é convidado a performar ações requisitadas pelo remetente, é possível perceber a presença da força ilocucionária que assinala as formas do modo imperativo.

Ao analisar o imperativo em termos mórficos, entende-se que na 2SG esse modo verbal apresenta formas verdadeiras (*deixa/recebe/abre/dá/diz/vai*) e supletivas (*deixe/receba/abra/dê/diga/vá*), alinhadas, cf. a norma gramatical (BECHARA, 2009 [1961]; ROCHA LIMA, 2013 [1972]; CUNHA; CINTRA, 2007 [1985]), aos modos indicativo e subjuntivo, respectivamente. Nos excertos transcritos, são formas verdadeiras *aceita*, *desculpa*, em (01), e *lembra*, em (03), provenientes da 2SG do indicativo com perda do morfe número-pessoal “-s”. A forma *queira*, em (03), por outro lado, é supletiva, uma vez advinda da 2SG do subjuntivo sem alterações morfológicas.

No que concerne à sintaxe, com base em Scherre (2007) e Cardoso (2006), o imperativo de 2SG, expressa-se vinculado a um sujeito – nulo, como em (01) e (03), ou pleno, em (02) – que pode constituir uma forma pronominal *tu* ou *você*, referidos, nesta ordem, por *prestarás*, em (01), e por *você*, em (02) e (03), ou uma forma nominal de tratamento⁷. Ao conjugar as formas imperativas (indicativa ou subjuntiva) às referências de sujeito pronominal de 2SG,

6 Na seção final de cartas pessoais, é comum encontrar tradições discursivas (KABATEK, 2006) que manifestam o uso do imperativo por meio de pedidos de desculpas e abraços.

7 Na expressão do imperativo, nem sempre é possível vincular a forma verbal a um sujeito pronominal, pois o sujeito pode não se realizar foneticamente. Nesses casos, o imperativo de 2SG tem sua referência de sujeito ocupada por uma forma nominal de tratamento.

fica evidente a existência de três construções imperativas⁸ no PB cf. o quadro (01), adaptado de Rumeu & Carvalho (2018, p. 396), com os exemplos desta seção.

Quadro (01) – As construções imperativas de 2SG do PB

REPRESENTAÇÃO DO IMPERATIVO			
PRONOME-SUJEITO	FORMA VERBAL	CLASSIFICAÇÃO	EXEMPLO
<i>Tu</i>	Indicativo	verdadeiro	<i>aceita, desculpa</i> (tu)
<i>Você</i>	Subjuntivo	supletivo	<i>queira transmitir</i> (você)
<i>Você</i>	Indicativo	abrasileirado	<i>lembra</i> (você)

Fonte: Rumeu & Carvalho (2018, p. 396)

De acordo com o quadro (01), a construção imperativa é verdadeira, como no exemplo (01), com uma forma indicativa (*aceita, desculpa*) com referência de sujeito *tu*; supletiva, tal como no fragmento (02), quando relacionada ao sujeito *você* com forma subjuntiva (*queira*); e, por fim, *abrasileirada*, em contexto de *você-sujeito* e forma originada do indicativo (*lembra*). Essa última deve sua emergência ao surgimento do *você* no sistema pronominal (LOPES, 2007; LOPES & CAVALCANTE, 2011) que atingiu o imperativo de 2SG, possibilitando a alternância entre a forma verbal e opção de sujeito.

Em linhas gerais, o ato ilocucionário, por meio do qual o imperativo se evidencia com um pedido, uma ordem ou uma súplica, está pautado formalmente em aspectos morfossintáticos. Nesse caso, é relevante ressaltar a existência no PB de formas verdadeiras e supletivas alinhadas em distribuição não complementar aos contextos de sujeito *tu* e *você*. Assim, descrito o imperativo de 2SG em termos semânticos, morfológicos e sintáticos, passa-se à próxima seção com abordagem de alguns estudos sobre o paralelismo linguístico.

Algumas considerações sobre paralelismo linguístico na expressão variável do imperativo

O paralelismo consiste em uma tendência geral à ocorrência de formas gramaticais similares juntas, como um mecanismo de repetição em prol da harmonização nos textos (SCHIFFRIN, 1981; WEINER & LABOV, 1983; SCHERRE, 1988; 1998). Essa é uma tendência, percebida em níveis lexicais, fonológicos, morfológicos, sintáticos e semânticos da língua, constituindo uma variável capaz de influenciar fenômenos de variação e mudança linguística, como a expressão variável do imperativo de 2SG no PB. Neste estudo, esse condicionamento é abordado a partir de dois grupos de fatores: o paralelismo sintático e o paralelismo fônico, descritos na sequência.

O paralelismo sintático é interpretado como uma variável do plano sintagmático que promove a ocorrência de formas precedentes e subsequentes que sejam semelhantes entre si no nível da sentença. Nesse caso, uma vez que o imperativo verdadeiro está associado formalmente

⁸ As construções imperativas podem ser entendidas, com base em Goldberg (1995, 2006), como um pareamento forma e função com propriedades semânticas e sintáticas específicas. Uma análise mais sólida sobre o imperativo na perspectiva construcional pode ser lida em Carvalho (2021).

ao *tu-sujeito* e o imperativo supletivo, ao *você-sujeito*, aventa-se a possibilidade, tal como demonstrado por Cardoso (2009, p. 117) e Diniz (2018, p. 88), de que formas associadas ao paradigma de *tu* (*tu-sujeito, de ti, para ti, contigo, teu/tua, imperativo verdadeiro*) impulsionem o imperativo com forma indicativa (*deixa/recebe/abre/dá/diz/vai*) e formas associadas ao contexto de *você* (*você-sujeito, de você, para você, com você, seu/sua, imperativo supletivo*) deflagrem o imperativo relacionado ao subjuntivo (*deixa/recebe/abre/dá/diz/vai*).

Enquanto Cardoso (2009) analisa a distribuição das formas imperativas em relação a esse fator em uma amostra sincrônica com dados orais de um grupo de indivíduos nativos de Fortaleza e residentes do Distrito Federal, Diniz (2018) volta seu olhar ao passado, investigando a variação do imperativo de 2SG em cartas cariocas do século XIX e XX. Os resultados dessas pesquisas se encontram sumarizados na tabela (01).

Tabela (01) – O efeito do paralelismo sintático sobre as construções imperativas de 2SG

PARALELISMO SINTÁTICO		INDICATIVO		SUBJUNTIVO	
ESTUDOS	PARADIGMAS	OCORRÊNCIAS/(%)	PESOS RELATIVOS	OCORRÊNCIAS/(%)	PESOS RELATIVOS
Cardoso (2009)	<i>tu</i>	205/233 (88%)	0.71	27/233 (22%)	-
	<i>você</i>	24/106 (23%)	0.10	82/106 (77%)	-
Diniz (2018)	<i>tu</i>	223/362 (61%)	-	139/362 (39%)	-
	<i>você</i>	62/350 (18%)	-	288/350 (88%)	-

Fonte: dados de Cardoso (2009) e Diniz (2018)

Cardoso (2009, p. 117) atesta que a precedência do dado por forma indicativa favorece o imperativo verdadeiro (88%, 205 oco, 0.71), mas por forma subjuntiva não (22%, 27 oco, 0.10). No estudo de Diniz (2018, p. 87), embora não tenha sido selecionado como estatisticamente relevante pelo programa GoldVarbX, esse fator se mostrou produtivo, na medida em que as formas imperativas do subjuntivo, em sua maioria, foram precedidas por formas do paradigma de *você* (82%, 288 oco) e as formas imperativas do indicativo foram antecedidas majoritariamente por formas do paradigma de *tu* (61%, 223 oco). Ainda que os estudos adotem perspectivas temporais distintas, ambos comprovam a influência do paralelismo sintático no fenômeno, o que parece apontar para a potência desse fator ao longo do tempo.

O paralelismo fônico consiste em uma variável de atuação no plano da palavra (SCHERRE, 1998, p. 38) que permite a verificação do efeito de fonemas anteriores nos posteriores. Na expressão do imperativo de 2SG, esse fator é compreendido tendo em vista a *harmonização vocálica* entre a vogal precedente e a vogal final da forma imperativa, o que leva às expectativas de que

(a) verbos de paradigma regular de 1ª conjugação de paradigma mais marcado (com vogal precedente [+aberta] = *fala/fale, olha/olhe, espera/espere*) tendam a favorecer o imperativo verdadeiro (SCHERRE, 1998, p. 67; 2004, p. 18; 2007, p. 207);

(b) verbos de paradigma regular de 1ª conjugação de paradigma menos marcado (com vogal precedente [-aberta] = *vira/vire, use/usa, imagina/imagine*) tendam a disseminar o imperativo supletivo (SCHERRE, 1998, p. 67; 2004, p. 18; 2007, p. 207);

(c) verbos de paradigma regular de 2ª e 3ª conjugações (*come/coma, abre/abra*) tendam a impulsionar o imperativo supletivo (SCHERRE, 2004, p. 18; 2006, p. 13);

(d) verbos de paradigma irregular de oposição mais marcada (*faz/faça, traz/traga, diz/diga*) tendam a induzir o imperativo supletivo (SCHERRE, 2004, p. 18; 2006, p. 13; 2007, p. 207);

(e) verbos de paradigma irregular de oposição menos marcada (*dá/dê, vai/vá, sai/saia*) tendam a estimular o imperativo supletivo (SCHERRE, 2004, p. 18; 2006, p. 13); e de que

(f) verbos de paradigma especial de oposição menos marcada (*esquece/esqueça, corre/corra, segue/siga, sobe/suba, recebe/receba*) tendam a beneficiar as construções do imperativo verdadeiro (SCHERRE, 2004, p. 18; 2006, p. 13; 2007, p. 207).

As análises de Scherre (1998, 2004, 2006, 2007), baseadas em dados do imperativo de *corpora* contemporâneos, são comprovadas, em parte, por Diniz (2018, p. 145), em sua pesquisa com missivas históricas, cujos resultados encontram-se em síntese na tabela (02).

Tabela (02) – O efeito do paralelismo fônico sobre as construções imperativas de 2SG associadas ao indicativo nas cartas cariocas

PARALELISMO FÔNICO	CONSTRUÇÕES IMPERATIVAS DE 2SG (INDICATIVO)	
	OCORRÊNCIAS/(%)	PESOS RELATIVOS
Verbo regular de 1ª conj. com vogal precedente [+aberta]= {fala/fale, olha/olhe, espera/espere}	91/287 (32%)	0.484
Verbo regular de 1ª conj. com vogal precedente [-aberta]= {vira/vire, use/usa, imagina/imagine}	36/83 (43%)	0.420
Verbos regulares de 2ª e 3ª conjugações {come/coma, abre/abra}	06/18 (33%)	0.244
Verbos irregulares com oposição mais marcada {faz/faça, traz/traga, diz/diga, vê/veja, sê/seja}	78/176 (44%)	0.575
Verbos de paradigma irregular de oposição menos marcada (dá/dê, vai/vá, sai/saia)	12/42 (29%)	0.237
Paradigma especial {esqu[ε]ce / esqu[e]ça, c[ɔ]rre / c[o]rra}	67/126 (53%)	0.621
Valor de aplicação: indicativo Input: 0.365 Significance:0.050	290/732 (40%)	

Fonte: dados de Diniz (2018, p. 145)

Nas cartas, a linguista atesta que (a) verbos de paradigma regular de 1ª conjugação de paradigma mais marcado (0.484), (b) verbos de paradigma regular de 1ª conjugação de paradigma menos marcado (0.420), (c) verbos de paradigma regular de 2ª e 3ª conjugações (0.244) e (e) verbos de paradigma irregular de oposição menos marcada (0.237) não favorecem o imperativo verdadeiro; enquanto (d) verbos de paradigma irregular de oposição mais marcada (0.575) e (f) verbos de paradigma especial de oposição menos marcada (0.621) conduzem à ocorrência do imperativo com formas indicativas. Ainda que esses resultados difiram dos de Scherre (1998, 2004, 2006, 2007), é importante salientar que no estudo de Diniz (2018), o paralelismo fônico foi o 5º fator estatisticamente relevante na depreensão do estatuto do imperativo de 2SG, o que denota a importância dessa variável na explicação do fenômeno.

Em suma, o paralelismo linguístico (SCHIFFRIN, 1981; WEINER & LABOV, 1983; SCHERRE, 1988; 1998), concebido como a regularidade de ocorrência de formas similares na língua, tende a constituir um fator condicionante em fenômenos variáveis. Resta saber como essa variável, especificamente no plano sintático e no plano fonético, influencia a manifestação das formas imperativas nas missivas mineiras. Antes, porém, os procedimentos teórico-metodológicos adotados para a depreensão desse efeito serão apresentados na seção a seguir.

Fundamentos teórico-metodológicos: a Sociolinguística Histórica aplicada ao trabalho com as missivas mineiras

A Sociolinguística Histórica parte do Princípio do Uniformitarismo⁹ (LABOV, 1994, p. 21) para, em cotejo com o presente, investigar os fenômenos linguísticos em sincronias pretéritas. Labov (1972, p. 100) afirma que o trabalho do linguista histórico consiste em fazer “o melhor uso dos maus dados”¹⁰, maus no sentido de serem fragmentados, corrompidos e distantes do vernáculo. Para lidar com essa realidade desafiadora, Romaine (1982)¹¹, precursora dos estudos em Sociolinguística Histórica, propõe um aporte metodológico eficiente a ser empregado em amostras escritas seguindo o rigor necessário a pesquisas científicas. A linguista reforça a atenção a ser dada ao “problema dos filtros” (ROMAINE, 1982), na medida em que cabe ao pesquisador, em seu percurso ao “reino das traças” (LOBO, 2009, p. 307), estar atento a marcas que ora se aproximam, ora se distanciam do vernáculo. Nesse sentido, cabe abordar

9 O Princípio do Uniformitarismo (*The Uniformitarian Principle*, LABOV, 1994, p. 21) se baseia na ideia do “uso do presente para explicar o passado”, na medida em que os fatos linguísticos de sincronias pretéritas repercutem na realidade linguística da contemporaneidade.

10 The great art of the historical linguist is to make the best of this bad data, “bad” in the sense that it may be fragmentary, corrupted, or many times removed from the actual productions of native speakers (LABOV, 1972, p. 100).

11 Romaine (1982) realizou um estudo diacrônico da variação das cláusulas relativas na língua inglesa com base em fatores linguísticos e extralinguísticos, consolidando a emergência da Sociolinguística Histórica.

as questões de *autenticidade*, *autoria* e *validade social e histórica*¹², propostas por Hernández Campoy & Schilling (2012, p. 63-79), que foram empregadas nesta pesquisa.

A *autenticidade* se refere à dificuldade de “garimpar” o *corpus* em busca do vernáculo, já que os textos históricos, uma vez escritos, não refletiriam a utilização espontânea da língua. Neste estudo, a questão da autenticidade está assegurada pela escolha do gênero textual utilizado. As cartas pessoais, gênero com características relativamente estáveis (BAKHTIN, 2003 [1953]) que consiste no estabelecimento de um diálogo à distância, permitem acessar o vernáculo por meio da escrita (AGUILAR, 1998, p. 239, 240) graças à sua natureza dialógica que as aproxima da oralidade (MARCUSCHI, 2001).

A questão da *autoria* está relacionada à necessidade de certificar se a tinta que assina o manuscrito pertence de fato ao autor a quem é atribuída. Em outros termos, uma vez que no passado é comum encontrar textos apócrifos (redigidos por amanuenses), para o sociolinguista importa saber se está de fato diante de um texto realmente autógrafo (produzido pelo autor). Essa questão foi certificada pelo rigoroso aporte filológico (SPINA, 1977; ACIOLI, 1994; MEGALE & CAMBRAIA, 1999; CAMBRAIA, 2005) utilizado pelo grupo de pesquisa *Para uma sociolinguística Histórica do português brasileiro: variação sincrônica e mudança diacrônica*¹³, cujos integrantes produzem edições conservadoras que atestam a autoria dos testemunhos.

A *validade social e histórica* faz referência à obtenção de informações sobre a estrutura social de sincronias pretéritas, tendo em vista a dificuldade de resgatar dados sobre os escreventes da amostra que podem se perder ao longo do tempo. Neste trabalho, essa questão foi garantida pelo reconhecimento dos missivistas, em sua maioria ilustres, que assinam os manuscritos utilizados. Os perfis biográficos de mineiros, como Carlos Drummond de Andrade, João Pinheiro e Henriqueta Lisboa¹⁴, podem ser levantados de forma relativamente fácil nos acervos públicos¹⁵ (fontes primárias) em que essas cartas se encontram bem como em obras biobibliográficas (fontes secundárias) de escritores mineiros (DUARTE, 2018).

12 Hernández-Campoy & Schilling (2012, p. 63-79) abordam, além das questões de *autenticidade* (*authenticity*), *autoria* (*authorship*) e *validade social e histórica* (*social and historical validity*), as questões de *representatividade* (*representativeness*), *validade empírica* (*empirical validity*), *invariação* (*invariation*) e *ideologia padrão* (*standar ideology*). Nesta pesquisa as primeiras foram escolhidas em lugar das últimas por se mostrarem mais apropriadas às possibilidades da análise adotada neste estudo.

13 O projeto de pesquisa desenvolvido na Universidade Federal de Minas Gerais (UFMG) tem como intuito pesquisar fenômenos morfossintáticos do PB associados à reorganização do quadro pronominal com base na produção de edições de manuscritos oitocentistas e novecentistas disponíveis em acervos públicos na cidade de Belo Horizonte. O projeto, que se encontra em sua segunda fase, pode ser acessado pelo site http://www.letas.ufmg.br/sistemas/cpq/projeto_site.php?id=9.

14 Neste estudo, optou-se por preservar a identidade dos escritores das cartas, que ainda estão em análise. Nesse caso, a autoria será retomada com as letras iniciais dos seus nomes. Uma descrição mais completa da amostra pode ser vista em Carvalho (2020, p. 78).

15 As cartas utilizadas na pesquisa encontram disponíveis no Acervo dos Escritores Mineiros (AEM), no Arquivo Público Mineiro (APM), no Museu Abílio Barreto (MAB), no Instituto Histórico e Geográfico de Minas Gerais (IHGMG) e no Arquivo Público da Cidade de Belo Horizonte (APCBH).

Dessa maneira, ao lidar com as questões de *autenticidade, autoria e validade social e histórica* (HERNÁNDEZ-CAMPOY & SCHILLING, 2012), esta pesquisa busca atender aos critérios propostos pela Sociolinguística Histórica. Assim, munido dos instrumentos metodológicos necessários, o linguista pode proceder ao árduo trabalho nos acervos com os textos que resistiram à ação erosiva do tempo. Uma vez realizado o trabalho com as amostras em observância a esses critérios, passa-se ao processo de codificação e quantificação dos dados encontrados.

A realização de análises multivariadas para depreender os fatores internos (linguísticos) e externos (extralinguísticos) que condicionam uma regra variável é um dos pilares da Sociolinguística Variacionista (WEINREICH; LABOV & HERZOG, 2006 [1968]; LABOV, 2008 [1972]). Nos estudos que seguem essa orientação, é comum utilizar programas computacionais que auxiliam o pesquisador na obtenção dos resultados para a devida descrição e análise. Nesta pesquisa os dados de imperativo de 2SG mapeados nas missivas mineiras foram codificados para tratamento pelo programa GoldVarbX (SANKOFF; TAGLIAMONTE & SMITH, 2005)¹⁶, um confiável programa utilizado em pesquisas linguísticas de natureza quantitativa.

Pertencente ao pacote de programas VARBRUL (VARiAble RULE), o GoldVarbX é um programa de cômputo de variáveis adaptado ao sistema Windows desenvolvido por pesquisadores canadenses (GUY & ZILLES, 2007). Essa importante ferramenta computacional gera os índices de frequência de uma regra variável em relação aos fatores linguísticos e extralinguísticos de um fenômeno, indicando, por meio dos pesos relativos, que condicionamentos atuariam de modo mais ou menos propulsor sobre essa variação. Esses índices são gerados basicamente por meio de duas rodadas: a *make cell* e a *binomial*. Na primeira, o programa gera a frequência de uso da variável dependente, em termos numéricos e percentuais, de maneira geral e em relação a cada grupo de fator utilizado para a pesquisa. Na segunda, a partir de um valor de aplicação da regra variável, o GoldVarbX apresenta, em pesos relativos, os fatores estatisticamente relevantes para aplicação da variável dependente.

Nesta pesquisa, a variável dependente consiste na expressão binária do imperativo de 2SG: imperativo verdadeiro (indicativo) vs. imperativo supletivo (subjuntivo). Os grupos de fatores utilizados na rodada *make cell*, com base em estudos linguísticos sobre a variação do imperativo (SCHERRE, 2003, 2007, 2012; CARDOSO, 2009; DINIZ, 2018; RUMEU, 2019) foram: *sujeito de 2SG; paralelismo sintático; tipo de conjugação do verbo; paralelismo fônico; tipo de pronome átono; número de sílabas do verbo; polaridade da estrutura; padrão sintático da sentença; tipo de verbo; período das cartas (1868-1993); subgênero da carta (amorosa, amistosa, familiar); gênero e faixa etária do missivista*. Desses grupos, tomando como valor de aplicação, o imperativo verdadeiro (indicativo), foram selecionados na rodada *binomial* (*Best stepping up run: #49; Log likelihood: -108.059; significance: 0.008*), como estatisticamente

16 O Goldvarb (Pacote de Programas VARBRUL – VARiAble RULE) está disponível para acesso no endereço eletrônico <http://individual.utoronto.ca/tagliamonte/goldvarb.html>.

relevantes, os grupos de fatores: *paralelismo sintático; o contexto de sujeito de 2SG; o subgênero da missiva; a polaridade da construção imperativa e o paralelismo fônico*, nessa ordem.

De todos esses grupos, cujos resultados podem ser averiguados com profundidade em Carvalho (2020), serão apresentadas a seguir, como recorte proposto neste artigo, as análises dos dados referentes ao paralelismo sintático e ao paralelismo fônico, logo após a discussão dos resultados gerais.

A disputa entre o indicativo e o subjuntivo nas cartas: resultados gerais

No *corpus* formado a partir de 202 missivas mineiras, por meio da rodada *makecell* no programa GoldvarbX, obteve-se a distribuição geral dos dados observada na tabela (03), que evidencia, como esperado, a sobreposição do imperativo supletivo sobre o imperativo verdadeiro.

Tabela (03) – Distribuição geral dos dados de imperativo de 2SG nas cartas mineiras

CONSTRUÇÕES IMPERATIVAS DE 2SG	
INDICATIVO (TU)	SUBJUNTIVO (VOCÊ)
73/388	315/388
(19%)	(81%)

Fonte: Carvalho (2020, p. 120)

Os dados da tabela (01) permitem evidenciar que foram encontradas 388 ocorrências de imperativo de 2SG na amostra, sendo 19% (73 oco) do imperativo verdadeiro (indicativo) e 81% (315 oco) do imperativo supletivo (subjuntivo). De (04) a (07), observam-se algumas ocorrências dessa variação no *corpus*.

(04) *Imagina, agora, Agenor em que condições vou ficar seduzido sem ter quem olhe para os meus pequenos e ainda com genros duvidosos! (RAAP. Lagoa Santa, 27.12.1913) – imperativo com forma indicativa*

(05) *Se soubesses a a afflicção com que espero o correio para ler as tuas cartas, me mandarias menos cartões! Mas as cartas assim pedidas não tem valor! Escreve mesmo os teos cartões! (JP. Rio de Janeiro, 15.02.1891) – imperativo com forma indicativa*

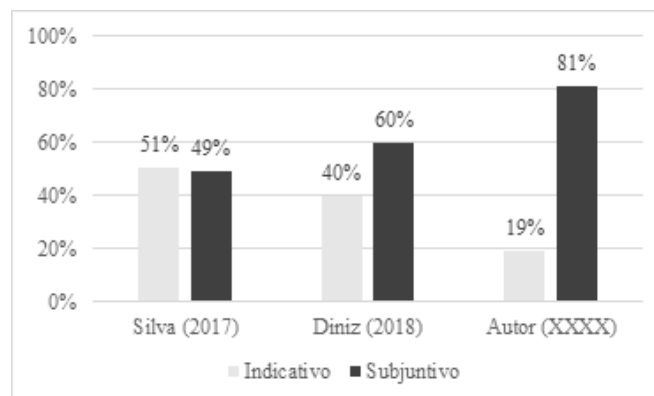
(06) *Faça de conta que debaixo de sua janela, da calçada de sua rua, sobe para você este canto de amizade, com os melhores desejos de paz, saude e vida, junto dos seus. (RCAM. Belo Horizonte, 20.10.1977) – imperativo com forma subjuntiva*

(07) *Você sabe o que é uma vida inteiramente [...] à burocracia? É ridículo mas trágico: afinal, tem uma certa grandeza [...]. Abrace este seu velho e grato Carlos (CDA. Rio de Janeiro, 01.06.1938) – imperativo com forma subjuntiva*

As formas imperativas nas cartas mineiras encontram-se em variação assim como nas cartas cariocas dos séculos XIX e XX analisadas por Silva (2017) e Diniz (2018). Os índices

percentuais desses trabalhos se diferem entre si como pode ser observado no gráfico (01), cujos dados possibilitam estabelecer algumas comparações.

Gráfico (01) – Distribuição geral das formas imperativas por estudo sociolinguístico



Fonte: dados de Silva (2017), Diniz (2018) e Carvalho (2020)

Os dados do gráfico (03) mostram que a disputa entre o indicativo e o subjuntivo na composição do imperativo de 2SG se difere substancialmente em relação ao *corpus*. Embora não seja o foco da análise aqui proposta, chama a atenção que, entre os mineiros, as formas do imperativo supletivo superem as do imperativo verdadeiro em uma dimensão proporcionalmente superior à encontrada entre os cariocas. Em Diniz (2018, p. 80), há uma frequência levemente maior do imperativo supletivo (indicativo: 40%, 290 oco; subjuntivo: 60%, 442 oco) e, em Silva (2017, p. 58), a variação entre as duas formas se mostra muito mais acirrada (50,8%, 400 ocorrências; subjuntivo: 49,2%, 387 ocorrências). Uma possível explicação para o evidente predomínio do imperativo com forma subjuntiva nas missivas mineiras estaria na prevalência de ocorrências imperativas em cartas de *você-sujeito* (70%, 191 cartas). Nesse sentido, uma vez o inovador *você* avançando progressivamente sobre os espaços funcionais do *tu*, sobretudo a partir da década de 1930 (RUMEU, 2013; 2016; 2019; DINIZ & RUMEU, 2019), essa parece ter sido a opção predileta de sujeito dos escreventes, corroborada por estudos da sincronia atual que alocam Minas Gerais nos domínios do *você* (SCHERRE, 2007, p. 192; LOPES & CAVALCANTE, 2011, p. 39). Assim, tendo em vista que o *você-sujeito* aciona o imperativo supletivo na tradição gramatical (BECHARA, 2009 [1961]; ROCHA LIMA, 2013 [1972]; CUNHA & CINTRA, 2007 [1985]), seria possível afirmar que os mineiros, legítimos representantes dessa tradição, tenham sido influenciados pela norma-padrão ao se expressarem no modo imperativo. Essa atuação, entretanto, não foi determinante ao ponto de impedir a variação, já que nem sempre a distribuição complementar entre sujeito pronominal e forma imperativa se manifestou, fato observado nas 30 ocorrências de *imperativo abrasileirado* (forma indicativa em contexto de *você-sujeito*, cf. PAREDES SILVA *et al.* 2000, p 121) no *corpus*.

Analisada a distribuição geral dos dados, parte-se para o efeito do paralelismo linguístico sobre expressão variável.

O paralelismo sintático no imperativo de 2SG nas cartas mineiras

Os dados referentes à atuação do paralelismo sintático na distribuição das formas imperativas (indicativa vs. subjuntiva) podem ser examinados na tabela (04), com os contextos controlados relacionados a esse grupo de fator.

Tabela (04) – O efeito do paralelismo sintático sobre as construções imperativas de 2SG associadas ao indicativo nas cartas mineiras

PARALELISMO SINTÁTICO		CONSTRUÇÕES IMPERATIVAS DE 2SG (INDICATIVO)	
		OCORRÊNCIAS/(%)	PESOS RELATIVOS
1 ^a OCORRÊNCIA	Primeira ocorrência	4/23 (17%)	0.148
	PARADIGMA DO <i>Você</i> (2SG)	Precedida por <i> você-sujeito pleno</i>	11/65 (7%)
	Precedida por <i> você não-sujeito (de você, para você, com você)</i>	-	-
	Precedida por formas pronominais de 3SG (<i>se, o/a, lhe, seu/sua</i>)	3/80 (4%)	0.196
	Precedida por formas imperativas no subjuntivo	8/102 (8%)	0.395
PARADIGMA do <i>Tu</i> (2SG)	Precedida por <i>tu-sujeito</i> (nulo ou pleno)	11/16 (69%)	0.751
	Precedido por <i>tu não-sujeito</i> (<i>de ti, para ti, contigo</i>)	2/3 (67%)	0.788
	Precedida por formas pronominais de 2SG (<i>teu/tua</i>)	22/41 (54%)	0.810
	Precedida por formas imperativas no indicativo	9/15 (60%)	0.814
PARADIGMA DA 2PL	Precedida por formas do paradigma de <i>vocês</i> (<i>seus, lhes, de vocês</i>)	2/7 (29%)	0.837
	Precedida por formas do paradigma de <i>vós</i> (<i>vos, vossos, vossas, convosco</i>)	1/9 (11%)	0.387
TOTAL		73/388 (19%)	

Fonte: dados do Autor

De maneira geral, os dados confirmam as conjecturas com base em Cardoso (2009, p. 116) e Diniz (2018, p. 88) sobre a variação do imperativo em função do paralelismo sintático,

na medida em que o paradigma de *tu* (*tu-sujeito, de ti, para ti, contigo, teu/tua, imperativo verdadeiro*) propulsiona o imperativo verdadeiro (0.751, 0.788, 0.810, 0.814) e o paradigma de *você* (*você-sujeito, de você, para você, com você, se, o/a, lhe, seu/sua, imperativo supletivo*) desencadeia o imperativo supletivo (0.694, 0.196, 0.395). Entre esses pesos relativos, salta aos olhos, entretanto, o fato de o *você* sujeito-pleno (0.694) mostrar-se como um contexto deflagrador da forma indicativa, o que pode ser lido como uma possível evidência não só da produtividade dessa forma pronominal no PB mineiro, mas também da existência de vestígios históricos do *imperativo abrasileirado* na amostra (30 oco). Por fim, vale mencionar o alto peso relativo associado às formas imperativas com indicativo precedidas pelo paradigma de *vocês* (0.837) que contrasta com a baixa probabilidade de expressão dessas formas em primeira ocorrência (0.148) e em contexto precedido pelas formas associadas ao *vós* (0.387).

De (08) a (18) estão dispostos alguns exemplos dessa variação em função de cada um desses fatores. Nesses fragmentos, fica evidente a pressão do paralelismo sintático sobre a expressão das formas verdadeiras e supletivas nos paradigmas de *tu* e *você*, respectivamente.

(08) *Avis*e ao Delzo que, sob registro nº 29.018, mande*ia* casimira hontem, seguindo pelo Bento 300\$000. (AR. Belo Horizonte, 05.11.1937) – *primeira ocorrência da forma imperativa*

(09) [...] Você deseja *lhe* de arranjar-se mais qualquer para o envio. Não *deixe* de ir ver o Axel Munthe. (AR. Belo Horizonte, 06.01.1937) – *forma imperativa precedida por você-sujeito pleno*

(10) Desejo saber, tia, se quer encomendar alguma casa do Chile para você. *Mande* dizer-nos. (AVP. Santiago, 06.11.196) – *forma imperativa precedida por você em outras funções sintáticas*

(11) [...] se lhe parecer razoável, *junte* ao cartão uma palavrinha sua a êle, que queria muito a Papai e que sempre me pediu notícias dele e suas. (JCL. Rio de Janeiro, 09.10.1947) – *forma imperativa precedida por formas de 3ª pessoa do singular (se, o/a, lhe, seu/sua)*

(12) Não deixe de ir ver o Axel Munthe. *Traga* autographos. Não deixe tambem de ir à França. (AR. Belo Horizonte, 06.01.1937) – *forma imperativa precedida por imperativo supletivo*

(13) Dize ao [...] que estou á espera dos cobres. Não *mandes* a carta do Cashley, que já paguei. (AR. Belo Horizonte, 15.01.1930) – *forma imperativa precedida por tu-sujeito nulo ou pleno*

(14) Tudo que eu digo para a minha negra também pertence a elle e por isso lembrando a todo o momento de ti, minha Helena, eu lembro do nosso bemsinho. *Manda* o Raymundo despachar o meu clach como encomenda pela estrada de ferro; veio a casa-ca e elle não veio elle. (JP. Ouro Preto, 09.11.1890) – *forma imperativa precedida por tu em outras funções sintáticas (de ti, para ti, contigo)*

(15) Lucia: como vae a tua alergia? Tive urticaria quasi o anno inteiro. *Consóla-te* commigo. Agora, só com reza brava. Venceremos! (AM. s/ local, s/ data) – *forma imperativa precedida por formas de 2ª pessoa do singular (te, teu/tua)*

(16) Não fôras tu, minha terna compa-nheira e a vida para mim seria detestavel! Ah! deixa, minha Helena, *deixa* que nestas paginas eu fale esta linguagem cheia d. sentimento [...]. (JP. Rio de Janeiro, 14.02.1891) – *forma imperativa precedida por imperativo verdadeiro*

(17) Selma e todas as meninas se [...] a vocês. Com o Antonio e a sobrinha, *receba* um abraço fraternal do Anibal. (AM. Rio de Janeiro, 01.01.1941) – *forma imperativa precedida por 2PL vocês (seus, lhes, de vocês – possessivos, clíticos, sintagmas complementos)*

(18) não julge. Vossa merce enterecei-ras estas exprecções, a franqueza hé quem as dita; *Envieme* outra de amizade e benevolencia e todos os meos dezejós ficarão satisfeitos. (JP. Ouro Preto, 21.12.1869) – *forma imperativa precedida por formas de 2PL vós (vos, vossos, vossas, convosco)*

O paralelismo fônico no imperativo de 2SG nas cartas mineiras

A influência do paralelismo fônico na variação das formas verdadeiras e supletivas do imperativo pode ser analisada na tabela (05), com os condicionadores aventados para esse grupo.

Tabela (05) – O efeito do paralelismo fônico sobre as construções imperativas de 2SG associadas ao indicativo nas cartas mineiras

PARALELISMO FÔNICO	CONSTRUÇÕES IMPERATIVAS DE 2SG (INDICATIVO)	
	OCORRÊNCIAS/(%)	PESOS RELATIVOS
Verbo regular de 1ª conj. com vogal precedente [+aberta]= {fala/fale, olha/olhe, espera/espere}	10/40 (25%)	0.564
Verbo regular de 1ª conj. com vogal precedente [-aberta]= {vira/vire, use/usa, imagina/imagine}	42/163 (26%)	0.647
Verbos regulares de 2ª e 3ª conjugações {come/coma, abre/abra}	1/10 (10%)	0.377
Verbos irregulares com oposição mais marcada {faz/faça, traz/traga, diz/diga, vê/veja, sê/seja}	10/104 (10%)	0.249
Paradigma especial {esqu[ε]ce / esqu[e]ça, c[ɔ]rre / c[o]rra}	10/56 (18%)	0.549
TOTAL	73/388 (19%)	

Fonte: dados do Autor

Considerando os resultados de Scherre (1998, 2004, 2006, 2007) e Diniz (2018), os dados confirmam, ao menos na maioria dos casos, as expectativas sobre a variação do fenômeno em função da regularidade dos fonemas nas formas imperativas. De fato, o imperativo verdadeiro tende a ser impulsionado por verbos regulares de 1ª conjugação com vogal precedente [+aberta] (0.564) e por verbos de paradigma especial (0.549). Também, como esperado, a ocorrência da forma indicativa é definitivamente desfavorecida com verbos regulares de 2ª e 3ª conjugações (0.377) e com verbos irregulares com oposição mais marcada (0.249). Todavia, chama a atenção o fato de que verbos regulares de 1ª conjugação com vogal precedente [-aberta] (0.647) constituam o contexto mais provável de ocorrência do imperativo verdadeiro, efeito que destoa do encontrado nos trabalhos anteriores (SCHERRE, 1998, 2004, 2006, 2007; DINIZ, 2018). Esse fato representa uma contravenção ao paralelismo fônico, tendo em vista que, por esse princípio, a vogal precedente dessas formas verbais [-aberta] tenderia a se harmonizar com a vogal final do imperativo supletivo, que é, por sua vez, fechada. Uma possível explicação para essa dissonância seria talvez a falta de equilíbrio numérico dos dados associados a cada fator dessa variável, situação muito comum em trabalhos que lidam com amostras históricas (LABOV, 1972; ROMAINE, 1982). A ver, portanto, se esse cenário se confirma em outros estudos com essa variável voltados para sincronias pretéritas.

De (19) a (24), cada fator é ilustrado com ocorrências de formas imperativas extraídas das cartas mineiras, confirmando a atuação do paralelismo fônico na expressão variável do imperativo de 2SG.

(19) Lucia: como vae a tua alergia? Tive urticaria quasi o anno inteiro. *Consóla*-te commigo. Agora, só com reza brava. Venceremos! (AM. s/ local, s/ data) – *forma imperativa com verbo de paradigma regular 1ª conjugação mais marcado (com vogal precedente [+aberta])*

(20) *Mande* um grande abraço a M. Celina, que esperamos em dezembro. (AM. Rio de Janeiro, 18.10.1946) – *forma imperativa com verbo de paradigma regular de 1ª conjugação menos marcado (verbo com vogal precedente [-aberta])*

(21) [...] portanto, *queira* enviar-me dois exemplares dessa obra, pelo reembolso postal [...]. (AGN. Silvestre Ferraz, 06.07.1949) – *forma imperativa com verbo de paradigma regular de 2ª e 3ª conjugações mais marcado*

(22) *Dê* notícias nossas aos daí. (JCL. Rio de Janeiro, 07.04.1945) – *forma imperativa como verbo de paradigma irregular com oposição menos marcada (dá/dê, vai/vá, sai/saia)*

(23) Lucia, se estás disposta a não ter conforto, *vem* desde já [...]. (AM. Rio de Janeiro, XX.01.1944) – *forma imperativa com verbo de paradigma irregular com oposição mais marcada (faz/faça, traz/traga, diz/diga)*

(24) Saudade as meninas e com o nosso querido Antonio. *receba* o abraço fraternal do Anibal (AM. Rio de Janeiro, 13.12.1945) – *forma imperativa com verbo de paradigma especial de oposição menos marcada (esquece/esqueça, corre/corra, segue/siga, sobe/suba, recebe/receba)*

Assim, tendo analisado o efeito do paralelismo (sintático e fonético) na expressão do imperativo de 2SG pelas mãos dos missivistas mineiros, passa-se, enfim, às palavras finais deste artigo.

Considerações finais

Neste estudo, investigou-se a disputa entre o indicativo (*deixa/recebe/abra/dá/diz/vai*) e o subjuntivo (*deixe/receba/abra/dê/diga/vá*) na constituição do imperativo de 2SG no PB, voltando o olhar para o efeito do paralelismo linguístico (SCHIFFRIN, 1981; WEINER & LABOV, 1983; SCHERRE, 1988; 1998) nessa variação. Com esse intuito, as formas imperativas encontradas em missivas mineiras oitocentistas e novecentista foram examinadas através do GoldVarbX (SANKOFF; TAGLIAMONTE & SMITH, 2005) a fim de depreender a distribuição geral da variável dependente bem como a atuação do paralelismo sintático e do paralelismo fônico nessa disposição. Assim, visando tecer algumas generalizações sobre esta investigação, retomam-se as questões propostas nas considerações iniciais.

(a) Considerando os estudos relacionados ao fenômeno variável em uma perspectiva histórica (SILVA, 2017; DINIZ, 2018; RUMEU, 2019), as formas imperativas presentes no *corpus* seriam predominantemente alinhadas indicativo (imperativo verdadeiro) ou ao subjuntivo (imperativo supletivo)?

Na pesquisa com 202 cartas mineiras autógrafas disponíveis em acervos públicos, foram identificadas 388 ocorrências de imperativo de 2SG, sendo 73 (19%) de imperativo verdadeiro e 315 (81%) de imperativo supletivo, alinhando-se aos estudos de Silva (2017) e Diniz (2018) com amostras históricas, em que as formas subjuntivas prevaleceram sobre as indicativas. Esse resultado parece estar ligado à influência da norma-padrão (BECHARA, 2009 [1961]; ROCHA LIMA, 2013 [1972]; CUNHA; CINTRA, 2007 [1985]) sobre os escreventes, que optaram pelo inovador *você* (LOPES, 2007; LOPES & CAVALCANTE, 2011) em vez do *tu*, repercutindo, na maioria dos casos, o imperativo supletivo. Essa força, entretanto, não foi determinante na medida em que, rompendo a distribuição complementar, rastros do *imperativo abraileirado* (30 oco) foram encontrados na amostra.

(b) Considerando que o fenômeno é motivado por fatores linguísticos e extralinguísticos (SCHERRE, 2007; RUMEU, 2019; CARVALHO, 2020), em que medida o paralelismo sintático e o paralelismo fônico (SCHERRE, 1998; 2007; CARDOSO, 2009; DINIZ, 2018) constituiriam variáveis significativas na expressão do imperativo de 2SG nas cartas mineiras?

Tanto o paralelismo sintático quanto o paralelismo fonético foram selecionados como *Diadorim*, Rio de Janeiro, vol. 24, número 1, p. 180 - 202, 2022.

estatisticamente relevantes (*significância*: 0.008) pelo programa GoldvarbX (SANKOFF; TAGLIAMONTE & SMITH, 2005) para depreensão da variável dependente (indicativo vs. subjuntivo), tal como atestado por Scherre (1998, 2007), Cardoso (2009) e Diniz (2018). Embora nem todos os fatores de cada grupo tenham impulsionado as formas imperativas esperadas segundo trabalhos sobre o tema, os resultados apresentados fortalecem a tese de que formas similares tendem a ocorrer juntas (SCHIFFRIN, 1981; WEINER & LABOV, 1983; SCHERRE, 1988; 1998). Essa correlação pôde ser analisada – e comprovada – neste estudo tanto no plano sintagmático da sentença quanto no plano fonético da palavra.

Em suma, a manifestação do imperativo de 2SG constitui um fenômeno instigante sobre qual existem muitos estudos ainda por se realizar. Um exemplo está na busca por motivações que expliquem as proporções distintas da distribuição das formas imperativas entre os escreventes cariocas (SILVA, 2017; DINIZ, 2018) e mineiros, que merece análises futuras de modo aprofundado. Outro se encontra no porquê de os verbos regulares com vogal precedente [-aberta] terem rompido o paralelismo fônico nas cartas mineiras, que também carece de estudos posteriores. Com esse trabalho, ao investigar a variação do imperativo de 2SG em sincronias passadas, buscou-se jogar mais luz sobre o caminho para realização de pesquisas sociolinguísticas que vão do presente ao passado para o entendimento dos fenômenos que confirmam o caráter heterogêneo e ordenado das línguas naturais.

Referências

ACIOLI, V. L. C. *A escrita no Brasil Colônia: um guia para a leitura de documentos manuscritos*. Recife: UFP/Fundação Joaquim Nabuco/Massangana, 1994.

AGUILLAR, R. C. Presencia de lo oral en lo escrito: la transcripción de las declaraciones en documentos indianos del siglo XVI. In: OESTERREICHER, W.; STOLL, E.; WESCH, A. (ed.) *Competencia escrita, tradiciones discursivas y variedades lingüísticas: aspectos del español europeo y americano en los siglos XVI y XVII*. Tübingen: Narr, 1998.

BECHARA, E. *Moderna Gramática Portuguesa*. Rio de Janeiro: Lucerna, 2009 [1961].

BAKHTIN, M. Os gêneros do discurso. In: BAKHTIN, M. *Estética da criação verbal*. São Paulo: Martins Fontes, 2003 [1953]. p. 261-306.

CAMBRAIA, C. N. *Introdução à crítica textual*. São Paulo: Martins Fontes, 2005.

CARDOSO, D. B. B. *Variação e Mudança no Português Brasileiro: gênero e identidade*. Tese (Doutorado apresentada ao Programa de Pós-Graduação da Universidade de Brasília) – UNB, Brasília, 2009.

CARVALHO, L. F. *O estatuto variável do imperativo de 2ª pessoa do singular em missivas mineiras: um estudo sociolinguístico de cunho histórico (séculos XIX e XX)*. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2020.

Diadorim, Rio de Janeiro, vol. 24, número 1, p. 180 - 202, 2022.

CARVALHO, L. F. de. A expressão variável do imperativo no português brasileiro: uma análise sob o viés construcional. *Domínios de Linguagem*, [S. l.], v. 15, n. 4, p. 1022–1058, 2021. DOI: 10.14393/DL48-v15n4a2021-5.

CONDE SILVESTRE, J. C. *Sociolinguística histórica*. Madrid: Gredos. 2007.

CUNHA, C.; CINTRA, L. F. L. *Nova Gramática do Português Contemporâneo*. Rio de Janeiro: Nova Fronteira, [1985] 2007.

DINIZ, J. S. *A expressão variável do imperativo de 2ª pessoa do singular no português brasileiro: análise de cartas pessoais dos séculos XIX e XX*. 2018. Dissertação (Mestrado em Estudos Linguísticos) – Universidade Federal de Minas Gerais, Belo Horizonte, 2018.

DUARTE, C. L. *Dicionário biobibliográfico de escritores mineiros*. Autêntica, 2018.

FARACO, C. A. *The Imperative Sentence in Portuguese: a semantic and historical discussion*. Tese (Doutorado em Linguística) – University of Salford, Salford, 1982.

FARIA, I. H. O uso da linguagem. In: MATEUS, M. H. M.; BRITO, A. M.; FARIA, I. H.; FROTA, S.; MATOS, G.; OLIVEIRA, F.; VIGÁRIO, M.; VILLALVA, M. *Gramática da língua portuguesa*. Lisboa: Caminho, 2006. p. 55-84.

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

GOLDBERG, A. E. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press, 2006.

GUY, G.; ZILLES, A. *Sociolinguística Quantitativa – instrumental de análise*. São Paulo: Parábola Editorial, 2007.

HERNÁNDEZ-CAMPOY, J. M.; CONDE-SILVESTRE, J. C. *The Handbook of Historical Sociolinguistics*. Wiley-Blackwell, 2012.

HERNÁNDEZ-CAMPOY, J. M.; SCHILLING, N. The Application of the Quantitative Paradigm to Historical Sociolinguistics: Problems with the Generalizability Principle. In: HERNÁNDEZ-CAMPOY, J. M.; CONDE-SILVESTRE, J. C. *The Handbook of Historical Sociolinguistics*. Wiley-Blackwell, 2012, p. 63-79.

KABATEK, J.. Tradições discursivas e mudança linguística. In: LOBO, T. C. F.; RIBEIRO, I. M. O.; CARNEIRO, Z. O. N.; ALMEIDA, N. L. F. (ed.): *Para a história do português brasileiro: novos dados, novas análises*. Salvador: EDUFBA, 2006. p. 589-606.

LABOV, W. *Sociolinguistics Patterns*. Philadelphia: University of Pennsylvania Press, 2008 [1972].

LABOV, W. *Principles of Linguistic Change: Internal Factors*. Cambridge: Blackwell Publishers, 1994, v. I.

LOBO, T. C. F. Arquivos, acervos e reconstrução histórica do português brasileiro. In: OLIVEIRA, K.; CUNHA E SOUZA, H. F.; SOLEDADE, J. (org.) *Do português arcaico ao português brasileiro: outras histórias*. Salvador: EDUFUBA, 2009.

LOPES, C. R. S. Pronomes pessoais. In: BRANDÃO, S F & VIEIRA, S. R. (org). *Ensino de gramática: descrição e uso*. São Paulo: Contexto, 2007, p. 103-114.

LOPES, C. R. S.; CAVALCANTE, S. O. A cronologia do voçamento no português brasileiro: expansão de você-sujeito e retenção do clítico-te. In: *Linguística*. 2011, v.25, p. 30-65.

MARCUSCHI, L. A. *Da fala para a escrita: atividades de retextualização*. São Paulo: Cortez, 2001.

MEGALE, H.; CAMBRAIA, C. N. Filologia portuguesa no Brasil. *DELTA*, v. 15, n.spe, p. 1-22, 1999.

PAREDES SILVA, V. L.; SANTOS, G.; RIBEIRO, T. Variação na 2ª pessoa: o Pronome sujeito e a forma do imperativo. *Revista Gragoatá*. UFF, v. 9, n. 9, p. 115-123, 2000.

ROCHA LIMA, C. H. *Gramática Normativa da Língua Portuguesa*, Rio de Janeiro, José Olympio, 2013 [1972].

ROMAINE, S. *Socio-historical linguistics: its status and methodology*. Cambridge University Press. New York. 2010 [1982]. 159

RUMEU, M. C. B. *Língua e sociedade: a história do pronome você no português brasileiro*. Rio de Janeiro: Ítaca (FAPERJ), 2013.

RUMEU, M. C. B. Formas variantes do imperativo de segunda pessoa nos séculos XIX e XX: a expressão do social. *Signum: Estudos da Linguagem*, v. 19, p. 310-341, 2016.

RUMEU, M. C. B. A inserção do você no português brasileiro escrito dos séculos XIX e XX: reflexos nas construções imperativas de 2SG. *LaborHistórico*, v. 5, n. Especial 1, p. 15-38, 2019.

RUMEU, M. C. B.; CARVALHO, L. F. O imperativo em livros didáticos de língua portuguesa: a distância entre pesquisa e ensino. *Matraga*, v. 25, n. 44, p. 391-409, 2018.

SANKOFF, D.; TAGLIAMONTE, S. A.; SMITH, E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SCHERRE, M. M. P. *Reanálise da concordância nominal em português*. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.

SCHERRE, M. M. P. Paralelismo linguístico. *Revista de Estudos da Linguagem*. Belo Horizonte, v. 7, n. 2, p. 29-59, 1998.

SCHERRE, M. M. P. Norma e uso na expressão do imperativo em revistas em quadrinhos da Turma da Mônica. In: SILVA, D. E. G.; LARA, G. M. P.; MENEGAZZO, M. A. (org.). *Estudos de Linguagem: inter-relações e Perspectivas*. Campo Grande: Editora UFMS, 2003. p. 177-191.

SCHERRE, M. M. P. Norma e uso - o imperativo no português brasileiro. In: DIETRICH, W.; NOLL, V. (org.). *O Português do Brasil - Perspectivas da pesquisa atual*. Madrid/Frankfurt am Main: Iberoamericana - Vervuert, 2004, v. 1, p. 231-260.

SCHERRE, M. M. P. O imperativo gramatical no português brasileiro: reflexo de mudança linguística na escrita de revistas em quadrinhos. In: VOTRE, S.; RONCARATI, C. (ed.). *Anthony Julius Naro e a linguística no Brasil: uma homenagem acadêmica*. Rio de Janeiro: Letras, 2006. p. 306-319.

SCHERRE, M. M. P. Aspectos sincrônicos e diacrônicos do imperativo gramatical no português brasileiro. *Alfa*. 2007, v.51, n. 1, p. 189-222.

SCHERRE, M. M. P. Padrões sociolinguísticos do português brasileiro: a importância da pesquisa variacionista. *Tabuleiro de Letras*. n. 4, p. 01-32, jun. 2012.

SCHIFFRIN, D. Tense variation in narrative. *Language*, LSA, 57(1):5-62, mar, 1981.

SEARLE, J. R. *Speech acts: An essay in the philosophy of language*. Cambridge university press, 1969.

SILVA, E. N. *Formas imperativas de segunda pessoa no português brasileiro*. Tese (Doutorado em Língua Portuguesa) Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Rio de Janeiro, 2017.

SPINA, S. *Introdução à edótica: crítica textual*. 2ª ed. São Paulo: Cultrix/Edusp. 1977.

WEINER, E. J.; LABOV, W. *Constraints on the agentless passive*. In: *Journal of Linguistics*, 19(1983):29-58.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução de Marcos Bagno. São Paulo: Parábola, 2006 [1968].



A CONCORDÂNCIA NOMINAL COMO UM MARCADOR ESTILÍSTICO DE CONSTRUÇÃO DA PERSONA ESTEREOTIPADA DO ACADÊMICO DA UFSC¹

NOMINAL AGREEMENT AS A MARKER STYLE OF CONSTRUCTION OF THE STEREOTYPED PERSONA OF THE UFSC ACADEMIC

Ariele Helena Holz Nunes²

Sabrina Vieira Teixeira³

RESUMO

Este estudo tem por objetivo demonstrar como a concordância nominal pode se transformar em um marcador estilístico estereotipado no meio acadêmico, o qual influencia a construção da persona universitária. Para tanto, assumimos o trabalho reconfigurado com entrevistas sociolinguísticas, cuja coleta de dados seguiu o modelo laboviano variacionista. Ao todo, quinze entrevistas sociolinguísticas compuseram a amostra que será analisada no decorrer deste artigo. As entrevistas foram realizadas com acadêmicos de diversos cursos da Universidade Federal de Santa Catarina (UFSC), estratificados por deslocamentos e nível de letramento em quatro células sociais. O protocolo de coleta de dados foi replicado da amostra Deslocamentos (2019), do banco de dados Falares Sergipanos (FREITAG, 2013), e segue as diretrizes do projeto. À luz de Labov (1963, 1966, 1972a, 2010), Scherre (1988, 1997), Eckert (2005, 2012), Camacho (2010), Freitag, Martins e Tavares (2012), Hora (2014), Severo (2014), entre outros autores; chegamos ao entendimento de que a concordância nominal constitui um importante marcador estilístico presente na identidade da persona universitária. A relevância do estudo se dá, em termos gerais, na escolha em abordar o fenômeno da concordância nominal pela perspectiva dos estudos de terceira onda em uma amostra coletada nos moldes labovianos variacionistas de primeira onda.

PALAVRAS-CHAVE: Entrevistas sociolinguísticas; Três ondas da sociolinguística; Estilo; Construção da persona; Concordância nominal.

1 Este artigo é fruto das discussões e atividades propostas na disciplina de Sociolinguística e Dialectologia, oferecida pelo Programa de Pós-Graduação em Linguística (PPGL/UFSC) no semestre 2019/2. Agradecemos aos professores Dra. Isabel de Oliveira e Silva Monguilhott, Dra. Izete Lehmkuhl Coelho, Dra. Loremi Loregian-Penkal e Dr. Marco Antonio Martins pela partilha de conhecimentos e orientações de leitura e escrita. Estendemos os nossos agradecimentos aos pareceristas anônimos pelas sugestões que contribuíram com a melhoria deste artigo. As autoras também agradecem ao Projeto VARSUL e ao Projeto *Falares Sergipanos* pelo apoio.

2 Doutoranda em Linguística pelo Programa de Pós-Graduação em Linguística da Universidade Federal de Santa Catarina (UFSC), holz.arielle@gmail.com. A autora agradece à CAPES (processo 88887.639125/2021-00) pelo apoio financeiro.

3 Graduada em Pedagogia pela Universidade do Estado de Santa Catarina (UDESC). E-mail: vieiratsabrina@gmail.com.

ABSTRACT

This study aims to demonstrate how the nominal agreement can become a stylistic marker that influences the construction of the university persona. We assume the reconfigured work with sociolinguistic interviews, whose data collection follows the variationist labovian model. In all, fifteen sociolinguistic interviews constitute the sample that will be analyzed in this article. The interviews were conducted with academics from several courses at the Federal University of Santa Catarina (UFSC), which were stratified by displacements and level of literacy into four social cells. The data collection protocol was replicated from the Displacements sample (2019) of the Falaes Sergipanos database (FREITAG, 2013), and follows the guidelines of the project. Through the theoretical basis of Labov (1963, 1966, 1972a, 2010), Scherre (1988, 1997), Eckert (2005, 2012), Camacho (2010), Hora (2014), Freitag, Martins and Tavares (2012), Severo (2014), among other authors; we identify that nominal agreement represents an important stylistic marker in the university persona's identity. The relevance of the study is in the choice to approach the phenomenon of nominal agreement from the perspective of third wave studies in a sample collected in the first wave labovian variationist molds.

KEYWORDS: Sociolinguistic interviews; Three waves of sociolinguistics; Style; Persona construction; Nominal agreement.

Introdução

O percurso histórico da Sociolinguística estabelece uma relação intrínseca com as pesquisas labovianas, tanto em linhas teóricas quanto em procedimentos metodológicos. Nesse sentido, pode-se dizer que as entrevistas sociolinguísticas, implantadas inicialmente por Labov (1972a), na tentativa de investigar o vernáculo dos indivíduos de dada comunidade de fala, são um grande marco para a Sociolinguística, configurando em uma das principais estratégias metodológicas do campo. A priori, as entrevistas sociolinguísticas devem ser utilizadas para investigar os usos linguísticos e os padrões de fala que identificam uma comunidade mais ampla, todavia, podem ser adaptadas e tomadas para outros fins a critério do pesquisador.

Almejando descrever a experiência com o manuseio de entrevistas sociolinguísticas, este estudo seleciona como fenômeno linguístico a ser analisado a concordância nominal. A discussão tem por objetivo demonstrar como a concordância nominal pode se transformar em um marcador estilístico estereotipado no meio acadêmico, o qual influencia a construção da persona universitária. Tendo em conta esses fatores, o foco de análise deste trabalho sai da comunidade de fala, em que se incorpora o social como o reflexo da estrutura linguística (LABOV, 1972a) e adentra nos usos identitários que os falantes fazem da língua, enfatizando as práticas sociais e linguísticas como aspectos que se interligam e assumem o social como fundante do linguístico, trazendo à tona a variação estilística.

As concepções de Eckert (2005, 2012) acerca das três ondas da Sociolinguística embasam este estudo. Assumimos como parâmetro, sobretudo, os estudos de terceira onda, em que o foco é o estilo. Em relação à metodologia, este estudo se baseia na coleta⁴ e análise de entrevistas

4 A coleta de dados foi realizada em 2019.

sociolinguísticas, realizadas em um movimento de replicação da amostra Deslocamentos (2019), do *Projeto Falares Sergipanos* (FREITAG, 2013), as quais seguem as diretrizes do projeto. Foram utilizadas como objeto de investigação estilística quinze entrevistas sociolinguísticas que atualmente constituem o acervo do Projeto VARSUL⁵ (Variação Linguística na Região Sul do Brasil). A amostra foi construída em parceria com outros pesquisadores e compartilhada para a discussão de outros fenômenos além da concordância nominal.

O interesse, portanto, em transformar este trabalho em um estudo de terceira onda vem da observação da expressiva marcação de concordância nominal na fala dos acadêmicos-informantes. Sabendo que pesquisas quantitativas⁶ de primeira onda sobre a concordância de número (cf. FONSECA; FRANCESCHINI; LOREGIAN-PENKAL, 2018; LUCCHESI; DÁLIA, 2020 para uma visão), que comumente abordam condicionadores linguísticos⁷, já são recorrentes na literatura, este artigo se debruçará sobre a noção de estilo como o reflexo de práticas identitárias dos universitários da comunidade de fala em que as entrevistas foram coletadas. A marcação não só é influenciada pelo nível de letramento que os sujeitos carregam, como está ligada ao perfil social desses indivíduos. Além de frequentarem a academia, alguns informantes estão imersos em cursos, por vezes, tidos como elitizados, o que pode dar margem para a construção de um estereótipo referente às práticas linguísticas e sociais que a persona universitária deve exercer.

Acresce mencionar que esta proposta de trabalho reconfigura o tratamento de entrevistas sociolinguísticas (cf. GÖRSKI; VALLE, 2014; VALLE; GÖRSKI, 2014 para uma abordagem), saindo da perspectiva laboviana de primeira onda para centralizar o estilo do falante e não mais os padrões de fala da comunidade, olhando qualitativamente para os dados encontrados na amostra. O embasamento teórico das análises propostas está em Labov (1963, 1966, 1972a, 2010), Scherre (1988, 1997), Eckert (2005, 2012), Camacho (2010), Freitag, Martins e Tavares (2012), Hora (2014), Severo (2014), entre outros autores. Desta forma, partimos da trajetória histórica dos estudos sociolinguísticos, apresentamos os postulados das três ondas da Sociolinguística (ECKERT, 2012) com enfoque especial na terceira onda e destacamos a importância da variação estilística como prática identitária dos sujeitos sociais, que se valem de posturas repetidas socialmente e linguisticamente para construir a sua identidade.

Esperamos que as discussões que sucedem esta seção sejam de grande relevância para outros pesquisadores que escolherem abordar o estilo a partir do recurso metodológico

5 As entrevistas realizadas estão disponíveis no VARSUL (UFSC).

6 O trabalho de Guy e Zilles (2007) traz uma abordagem da Sociolinguística Quantitativa.

7 Reiteramos que pesquisas de primeira onda também abordam condicionadores extralinguísticos, mas enfocam a correlação entre variáveis linguísticas e macrocategorias sociais (classe, sexo, idade, etnia etc.) a fim de encontrar padrões sociolinguísticos regulares (LABOV, 1972a).

entrevista sociolinguística. Para além do que foi comentado, privilegamos o estilo por julgarmos necessário discutir identidade (cf. OUSHIRO, 2019 para uma visão sobre conceitos de identidade e pesquisas sociolinguísticas) e práticas linguísticas que validam a construção de uma persona em dado contexto social.

Pressupostos teóricos: os caminhos trilhados

A consolidação da Sociolinguística como área dos estudos linguísticos se deu a partir da década de 1960, com o lançamento de “Fundamentos empíricos para uma teoria da mudança linguística” de Weinreich, Labov e Herzog, citados a seguir como WLH (2006 [1968]); e as pesquisas desenvolvidas individualmente por Labov (1963, 1966). Os autores lançaram mão de conceitos fundamentais para o estudo da linguística social, sendo o principal: a ideia de variação linguística, em que a língua é compreendida como um sistema heterogêneo, que varia de maneira ordenada e sistemática. Em oposição a Saussure, que interpretava a língua como um sistema homogêneo, WLH (2006 [1968]) entendem a língua como um reflexo da estrutura social, permitindo que ela mude e os indivíduos continuem a se entender.

Severo (2014) sinaliza que Labov (1972a) concebia a língua como um fato social, uma estrutura heterogênea, de modo que o seu caráter variável e mutável se vincula a fatores sociais, estilísticos e linguísticos. O autor apostava na correlação entre forças estruturais e linguísticas atuando no funcionamento da língua, as quais se apresentavam nas diferentes comunidades de fala. Entretanto, Labov (1972a) reconhecia o social como exterior à língua. O autor abordava a língua, a mudança e a variação em termos de comunidade de fala, colocando o falante em segundo plano: “[...] a língua é um padrão abstrato localizado na comunidade de fala e exterior ao indivíduo” (LABOV, 2010, p. 7). O autor mencionava ainda que a língua é um fato social que depende de outros fatos sociais para o seu pleno funcionamento: “Para Labov, trata-se de um social atravessado por uma perspectiva empírica que privilegia a dimensão estrutural, como os níveis fonológico e morfossintático” (SEVERO, 2014, p. 37).

De acordo com Camacho (2010), a maior crítica feita à metodologia laboviana variacionista quanto ao social diz respeito ao fato da exteriorização do falante, ao seu esvaziamento enquanto condutor do próprio discurso. Nesse sentido, coube à Eckert (2000, 2005) se inquietar e reconfigurar o conceito de variável linguística, atribuindo a esse fenômeno um lugar privilegiado para a construção do significado social da linguagem. Na expectativa de adentrar ao significado social nos estudos variacionistas, a autora projetou a trajetória da pesquisa sociolinguística dos últimos quarenta anos em uma teoria que estabelece três ciclos de práticas analíticas, popularmente conhecidos como as três ondas da Sociolinguística (ECKERT, 2012).

Seguindo Eckert (2005, 2012), a primeira onda é marcada por estudos de natureza quantitativa, o foco está em estabelecer relação entre variáveis linguísticas e categorias macrosociais primárias, como: sexo, idade, escolaridade, classe socioeconômica etc. Os estudos de primeira onda são baseados na realização de entrevistas sociolinguísticas, em que se constitui uma amostra homogênea. Estratificados em células sociais, os informantes são agrupados em categorias sociodemográficas em uma comunidade de fala, de modo que se investiga o uso de certo fenômeno linguístico e a forma como este uso da língua é avaliado pelos próprios falantes.

Um recurso importante do qual os estudos labovianos se valeram na primeira onda é o paradoxo do observador, em Labov (1972a) definido como o meio de investigação que busca identificar como os sujeitos falam quando não estão sendo sistematicamente observados. Para tanto, esse tipo de leva metodológica fez uso de entrevistas sociolinguísticas tradicionais, em que se promove uma situação agradável para o uso do vernáculo⁸ através de questionamentos sobre a experiência pessoal do falante, fazendo com que o monitoramento seja baixo pelo envolvimento com a situação discursiva. De acordo com o autor, uma das questões que mais funciona para o alcance do vernáculo está relacionada com a temática “risco de morte”.

Na visão de Eckert (2005, 2012) o interesse das pesquisas labovianas de primeira onda se concentram na captura do vernáculo. À medida que o vernáculo é a primeira produção linguística, não sofre influência da correção social e tende a não ser monitorado. Tendo em conta esses fatores, pode-se dizer que a variação não foi tratada como uma escolha do falante, mas como resultado de um monitoramento natural, que surge das situações discursivas em que o indivíduo está inserido no uso que faz da língua. Em linhas gerais, os estudos da primeira onda tinham como centro a variação da fala em contextos distintos, impulsionada por diferentes tipos de motivações.

A primeira onda viu a mudança linguística emergir de dentro do sistema linguístico e ser influenciada por aspectos extralinguísticos:

A primeira onda viu a mudança linguística como emergindo de pressões dentro do sistema linguístico, primeiro afetando a fala daqueles menos sujeitos à influência da linguagem padrão e se espalhando através de populações cada vez mais resistentes à mudança. Ao mesmo tempo, uma variedade de variáveis que não são mudanças em andamento são estratificadas como resultado de coisas como contato dialetal e resistência à padronização. A perspectiva da primeira onda sobre o significado baseava-se na hierarquia socioeconômica: as variáveis eram tomadas para marcar o status socioeconômico e as dinâmicas estilísticas e de gênero eram vistas como resultantes dos efeitos dessas categorias na orientação dos falantes para o lugar atribuído nesta hierarquia (ECKERT, 2012, p. 90, tradução das autoras).

8 O vernáculo é o uso da língua que sofre menor monitoramento, o uso espontâneo, utilizado quando se fala com os amigos, com a família, em momentos informais com menos atenção prestada à fala (LABOV, 1972a).

Ainda nas teorizações de Eckert (2005, 2012), cabe mencionar que os estudos de primeira onda representam mudanças sonoras em andamento. Todavia, a variação não é motivada apenas pela fala, mas pelos condicionadores sociais que a cercam. Os padrões de fala podem mudar ao longo da vida, podem ser mais monitorados na juventude e menos monitorados na velhice, indicando que fatores sociais podem interferir no uso que se faz da língua. Resumidamente, as pesquisas do primeiro ciclo adentraram em uma perspectiva ampla sobre as categorias sociais, não aprofundando a motivação dessas categorias, o seu funcionamento em dada comunidade e o impacto nos falantes.

A segunda onda, por sua vez, caracteriza-se pelas pesquisas de cunho etnográfico, em que se volta o olhar para comunidades menores, em um intervalo de tempo consideravelmente longo, para alcançar um retrato de variáveis linguísticas locais (ECKERT, 2005, 2012). Para Camacho (2010), a ideia de investir em estudos etnográficos está atrelada ao interesse de reconhecer práticas locais mais salientes nos contextos de comunidades menores, através da observação estendida da dinâmica social. Freitag, Martins e Tavares (2012) complementam que os estudos de segunda onda também são de natureza quantitativa, mas se valem da perspectiva etnográfica para compreender como as categorias macrosociais funcionam e são motivadas na dinâmica local de certa comunidade. Esta dinâmica local corresponde às práticas específicas com a linguagem, de modo que traços identitários de dado grupo se tornam mais salientes nesses contextos, revelando uma noção de pertencimento àquela comunidade de fala, àquele grupo de falantes.

Especificando, esses traços locais salientes estão estritamente ligados ao significado social (ECKERT, 2000), ao passo que podem ser tomados como positivos, ao serem incorporados e reconhecidos, ou negativos, quando marcam diferenças sólidas com outros grupos, outras comunidades. O panorama dos estudos etnográficos deixa claro que se algumas variantes são estigmatizadas em um plano geral de uso da língua, quando passam a ser integradas e associadas a valores e práticas locais há a possibilidade de se tornarem positivas, uma vez que fazem parte da dinâmica local, logo expressam pertencimento (ECKERT, 2005, 2012; CAMACHO, 2010). Nesse tipo de abordagem, o enfoque recai sobre os conceitos de comunidades de fala e de identidade de grupo (FREITAG; MARTINS; TAVARES, 2012).

Na visão de Eckert (2005, 2012), a segunda onda emergiu para mostrar o lado social do vernáculo em termos de entidade representativa de um grupo, colocando-o como uma expressão identitária local. Dentre os trabalhos deste ciclo, Eckert (2012) cita três exemplos: a) o estudo de Labov (1972b) acerca do inglês vernacular afro-americano (AAVE), que revela o uso de traços vernaculares por adolescentes para marcar pertencimento a um grupo de pares, de modo que tal uso indexicaliza status e o pertencimento àquela comunidade de prática; b) a pesquisa de Milroy (1980), em que o autor se debruça sobre a variação fonológica nas redes sociais da comunidade de Belfast. Trata-se de focar comunidades de classe operária e examinar a relação entre engajamento local e uso do vernáculo (CAMACHO, 2010; FREITAG; MARTINS;

TAVARES, 2012), provando que o uso de variantes vernaculares se associa à cultura local, e abre espaço para o entendimento acerca da densidade e da multiplicidade da rede de relações do falante; e c) o estudo da autoria de Eckert (2000) sobre a função das categorias *jokers* e *burnouts* na indexação de classe socioeconômica em grupos de adolescentes em uma escola na área suburbana de Detroit. A partir dos usos e recursos linguísticos dos adolescentes, a variação é estratificada por classes: *jokers* para a classe média e *burnouts* para a classe trabalhadora.

Em Eckert (2005, 2012) se encontra a concepção de que os estudos de primeira e segunda onda focam na estrutura linguística e explanam, em segundo plano, a interferência do social na estrutura. Já os estudos de terceira onda adentram a estrutura e almejam demonstrar o significado social da variação, sendo o falante o ponto chave para entender o funcionamento de categorias macrossociais em determinada comunidade:

Os estudos de primeira e segunda ondas, segundo Eckert (2012), têm como foco a descrição da estrutura – um retrato estático. Os estudos de terceira onda incorporam a dinamicidade da estrutura, ou seja, como a estrutura se molda no cotidiano, com os condicionamentos sociais impostos e as relações de poder estabelecidas atuando sobre ela (FREITAG; MARTINS; TAVARES, 2012, p. 922).

Conforme antecipado, os estudos de terceira onda saem da esfera da comunidade, amplas na primeira onda e restritas na segunda, para adentrarem na perspectiva do indivíduo, buscando compreender a interação do falante com o seu grupo de pertencimento. Em vista disso, a terceira onda incorpora o social como parte fundante da variação, situada no falante. Nas palavras de Camacho (2010, p. 158): “[...] centra o foco na variação vista não como o reflexo do lugar social num ponto da escala, mas como um recurso para a construção de significado social”. O autor ainda complementa alegando que Eckert (2005, 2012) enxergou nos estudos de terceira onda a necessidade de conectar as categorias macrossociais e o seu funcionamento nas comunidades com a experiência do falante, chegando ao conceito de comunidade de prática.

Nos estudos de terceira onda, o foco de análise passa a ser as comunidades de prática, ao invés das comunidades de fala. As comunidades de prática⁹ são constituídas por indivíduos que compartilham valores, perspectivas, conhecimentos e interagem entre si com os mesmos objetivos. Interagindo e ao mesmo tempo, diferenciando-se de outras comunidades pelas práticas que exercem, as comunidades de prática marcam uma construção social de indivíduos pelas suas identidades e funções na sociedade. As comunidades de prática são distintas, podem ser grandes ou pequenas, relacionadas ou difusas, além disso, elas podem nascer ou morrer, e aceitam a alternância de membros em seu interior (ECKERT; MCCONNELL-GINET, 2010). Do mesmo modo, a identidade dos membros se relaciona com a sua própria participação na comunidade de prática, fortalecendo a ideia de pertencimento:

9 O trabalho de Salomão-Conchalo (2015) é um exemplo de estudo etnográfico que identificou diferentes comunidades de prática em uma escola de São José do Rio Preto-SP.

Comunidade de prática é um conjunto de pessoas agregadas em razão do engajamento mútuo em um empreendimento comum. Modos de fazer coisas, modos de falar, crenças, valores, relações de poder – em resumo, práticas – emergem durante sua atividade conjunta em torno do empreendimento. Como construto social, uma comunidade de prática é diferente da noção tradicional de comunidade, sobretudo porque é definida simultaneamente pelos seus participantes e pela prática na qual eles se engajam. Na verdade, são as práticas da comunidade e a participação diferenciada de seus membros nessas práticas que estruturam socialmente a comunidade (ECKERT; MCCONNELL-GINET, 2010, p. 102).

O foco nas comunidades de prática representa, dentre tantas coisas, que o estilo e a expressão da identidade através da variação são os pontos-chaves dessa onda. O estilo passa a ser visto como a combinação de variáveis distintas para a criação de um modo específico de fala (HORA, 2014), o qual traz à tona traços identitários de pertencimento a um grupo, à uma comunidade de prática. Na terceira onda, a agentividade do falante frente à variação passa a ser analisada. Se outrora o falante era uma entidade à margem dos estudos de variação: “Tal foco possibilita-nos ver o indivíduo como agente articulador de uma variedade de formas de participação em múltiplas comunidades de prática” (ECKERT; MCCONNELL-GINET, 2010, p. 103).

A variação em nível de estilo pode ser entendida como uma necessidade de adaptação aos diferentes contextos. Em outras palavras, um único falante pode carregar estilos múltiplos e uma identidade multifacetada por ser um sujeito social que atua em várias instâncias¹⁰. O estilo se tornou parte fundante do falante porque permite que ele transite por diferentes situações discursivas sabendo se colocar socialmente e linguisticamente (cf. ECKERT, 2000, 2005 para uma visão de estilo como prática social). Nesse sentido, tanto os indivíduos como as comunidades de onde são membros se fazem e refazem continuamente, tendo uma identidade múltipla e mutável.

Camacho (2010) salienta que os estudos de terceira onda buscaram privilegiar o significado social da variação ao enxergar no estilo uma forma de demonstrar que o falante possui práticas específicas, que criam a sua identidade tanto social quanto linguística. Hora (2014, p. 29) escreve ainda que a língua deixou de refletir o social para criá-lo:

Quando pensamos sobre a relação entre variação e grupos sociais, geralmente não são identificadas variáveis individuais. O significado da variação está em seu papel na construção dos estilos, e estudar o papel da variação na prática estilística envolve não simplesmente localizar variáveis nos estilos, mas entender essa localização como uma parte integral da construção do significado social (HORA, 2014, p. 29).

10 Oushiro (2019, p. 306) menciona o conceito de identidades sociolinguísticas, discutindo que as identidades são múltiplas e plurais.

A terceira onda entende a variação como o reflexo de identidades e categorias macrosociais, que deixam de ser vistas no plano geral e passam a ser incorporadas como parte de uma prática linguística estilística, em que o falante se insere na paisagem social através de práticas estilísticas que constituem a sua persona. Eckert (2005, 2012) pontua que uma característica do discurso de certa comunidade pode se destacar e ser retirado do momento de fala para representar um traço icônico daquele grupo, o que pode indexar o pertencimento àquele grupo sem ao menos ser utilizado efetivamente. Em outras palavras, se um falante faz uso de certo traço, ele pode ser considerado como membro daquele grupo, porque indexou em sua prática linguística características icônicas.

Esses traços de pertencimento, aos olhos dos indivíduos que estão no exterior da comunidade, podem servir como características de construção de um estereótipo. Tais traços identitários e estilísticos que marcam estereótipos podem ser utilizados de forma pejorativa, para marcar qualidades, ou podem expressar distinção entre os indivíduos em termos de pertencimento e não pertencimento. O pertencimento a uma comunidade indexicaliza uma série de significados relacionados, que não se formam acidentalmente. A variação passa a ser estilística na terceira onda à medida que os indivíduos criam ou se apropriam de um estilo de discurso que combina e vai ao encontro do seu estilo de vida. Para Eckert (2005, 2012), nesse jogo, falantes pertencentes a grupos exteriores se constituem como estrangeiros, e falantes pertencentes a grupos locais, configuram-se como nativos.

Os estudos de terceira onda partem do estilo para explicar o caráter social atribuído à linguagem e à variação¹¹. Os recursos estilísticos utilizados pelos falantes são diferentes dependendo da situação em que eles circulam. Seguindo essa lógica, é importante lembrar que o foco está para o indivíduo e não para a comunidade porque: “[...] estilos surgem de posturas repetidamente tomadas¹²” (ECKERT, 2012, p. 96, tradução das autoras).

Os recursos estilísticos são simbólicos, uma vez que trazem um novo leque de possibilidades de atribuir significado social à variação. Diante da nova perspectiva, os falantes passaram a ser agentes estilísticos, que adaptam recursos linguísticos em seus projetos de vida, nas situações sócio-discursivas cotidianas. Eles adaptam o estilo como uma forma de autoafirmação, de indexar sua identidade. Na terceira onda o estilo assume papel fundamental para explicar o social e as potencialidades do falante como membros de uma comunidade que ora fazem uso das mesmas práticas linguísticas e se diferem de outras comunidades, e ora apresentam traços individuais distintos, promovendo uma heterogeneidade:

O conceito-chave para o processo de construção é o de prática estilística. Até aqui, nos estudos variacionistas, o estilo tem sido tratado como ajustes à (in)formalidade da situação mediante o uso de variáveis individuais. A face renovada de estilo o identifica com o modo como os falantes combinam

11 Görski, Coelho e Nunes de Souza (2014) para uma coletânea de trabalhos sobre variação estilística.

12 No original: “[...] styles emerge from repeatedly taken stances” (ECKERT, 2012, p. 96).

variáveis para criar modos distintivos de fala, que fornecem a chave para a construção da identidade. A identidade consiste, por sua vez, em tipos particulares explicitamente localizados na ordem social. Continuamente, os falantes atribuem significado social à variação de um modo consequente, situação que implica certo grau de agentividade (CAMACHO, 2010, p. 159).

A questão nevrálgica dos estudos de terceira onda são os estilos, os quais exprimem significado social e dizem muito sobre quem o falante é e de onde ele atua ao evidenciar a função social e suas posturas repetidas (ECKERT, 2012). O estilo pode ser interpretado como o conjunto de ações dos falantes, em que se combinam, quando se pensa no linguístico, variáveis que criam modos distintos de se utilizar a língua: “O outro lado do estilo é como os falantes combinam variáveis para criar formas distintivas de falar. Essas formas de falar são uma chave para a produção das *personas* e - tipos sociais particulares que se localizam de forma explícita na ordem social” (HORA, 2014, p. 29). Esses modos, ao evocarem recursos estilísticos específicos, constroem uma persona, ou seja, o estilo projeta a persona que o falante deseja e simultaneamente, revela a sua identidade apenas com práticas ora sociais ora linguísticas. A persona emerge dos comportamentos e traços identitários do falante, além de ser moldada pelos papéis sociais que eles ocupam e pelos ambientes em que realizam práticas.

Um dos recursos linguísticos que podem atuar na construção da persona e marcar a identidade do sujeito é o fenômeno da concordância nominal (cf. OUSHIRO, 2015; SALOMÃO-CONCHALO, 2015; CAMACHO; SALOMÃO-CONCHALO, 2016 para trabalhos que abordam a concordância nominal em uma perspectiva identitária). Quanto a sua realização, Scherre (1997, p. 182) aponta que a concordância é um fenômeno variável no português falado no Brasil, apresentando: “[...] tanto a preservação das marcas redundantes (variantes explícitas) quanto a perda das marcas redundantes (variantes zero), condicionadas por fatores linguísticos e não-linguísticos”. Como qualquer outro fenômeno, a concordância nominal é uma variável dependente, que pode ser analisada a partir de variáveis independentes, dentre elas o estilo.

Faz-se necessário enfatizar que a concordância nominal é um dos lugares da gramática (morfossintaxe) em que ocorre a variação (cf. SCHERRE, 1988; FERNANDES, 1996; CARVALHO, 1997; SCHERRE; NARO, 1997; BRANDÃO, 2011; LUCCHESI; DÁLIA, 2020), abrindo espaço para que variantes disputem pela expressão dessa variável, sendo elas: a marcação de concordância e a não marcação. Este fenômeno possibilita uma análise por diferentes vieses, desde discussões quantitativas variacionistas tradicionais de primeira onda até o debate sobre a variação estilística (cf. SALOMÃO-CONCHALO, 2015), típicas de terceira onda. Há uma recorrência histórica nos estudos de concordância de número em abordá-la através de condicionadores linguísticos como: saliência fônica, posição no SN, classe gramatical, contexto fonético/fonológico seguinte, função sintática no SN, traços mórficos e semânticos (SCHERRE, 1988), entre outros. Contudo, este estudo se debruçará sobre a concordância nominal na função de marcador estilístico, almejando desenvolver uma investigação sobre a

sua frequência de marcação na fala de acadêmicos de níveis de letramento distintos. Abordamos a realização dessa variável como um traço identitário desses sujeitos, que reflete na construção de uma persona universitária.

Metodologia

Nesta seção, apresentamos os procedimentos metodológicos nos quais se formatou a pesquisa, descrevemos a coleta e o tratamento dos dados, e o tipo de análise realizada sobre o fenômeno da concordância nominal. As pesquisas que seguem um viés sociolinguístico e variacionista têm como característica o uso de uma metodologia empírica, já que se utiliza de dados reais, coletados de falantes reais, em situações reais de uso.

Constituição da amostra

Este estudo replicou a amostra Deslocamentos (2019), do banco de dados *Falares Sergipanos* e é constituído nos mesmos moldes de Freitag (2013). A coleta piloto¹³ foi realizada no ano de 2019 e serviu como objeto de análise de fenômenos linguísticos para os pós-graduandos da disciplina de Sociolinguística e Dialectologia do Programa de Pós-Graduação em Linguística da UFSC. Atualmente a amostra faz parte do acervo do Projeto VARSUL.

A amostra de coleta desta pesquisa se deu a partir de entrevistas realizadas com 15 estudantes do campus Florianópolis da Universidade Federal de Santa Catarina (UFSC). Esses entrevistados eram oriundos de 4 deslocamentos, sendo eles: a) Deslocamento 1 - nascido e residente em Florianópolis; b) Deslocamento 2 - nascido na região metropolitana de Florianópolis; c) Deslocamento 3 - Catarinense de outras regiões do estado; e d) Deslocamento 4 - vindos de outros estados.

Além disso, os informantes foram estratificados por nível de letramento. Foram realizadas entrevistas com estudantes de início de curso e de final de curso a fim de reconhecer padrões de marcação de concordância em estudantes de diferentes níveis de letramento. Pelo fato da coleta ser estratificada por deslocamento e fases dos cursos, houve certa dificuldade de encontrar estudantes que se encaixavam nos parâmetros e/ou que tinham disponibilidade para mais ou menos 60 minutos de entrevista. Em função disso, não foi possível estabelecer uma quantidade igual de entrevistas para cada deslocamento, já que a proposta inicial era preencher cada célula social (deslocamentos) com a mesma quantidade de acadêmicos, sendo 4 estudantes de cada um, 2 de fases iniciais (de 1ª a 3ª fase) e 2 de fases finais (de 6ª a 10ª fase). Porém, com as possibilidades de trabalho que encontramos, a distribuição dos entrevistados em cada célula se deu conforme demonstrado no Quadro 1:

13 A replicação da coleta foi possível mediante a abertura de emenda no projeto Falares Sergipanos (FREITAG, 2013), que foi submetido à apreciação do Comitê de Ética da Universidade Federal de Sergipe (UFS), vinculado ao Sistema Nacional de Informações sobre Ética em Pesquisa - SISNEP, sob aprovação do CAAE N° 0386.0.107.000-11.

Quadro 1 – agrupamento dos informantes por célula social

DESLOCAMENTOS - DISPOSIÇÃO DE CURSOS E FASES							
DESLOCAMENTO 1		DESLOCAMENTO 2		DESLOCAMENTO 3		DESLOCAMENTO 4	
FASES INICIAIS	FASES FINAIS	FASES INICIAIS	FASES FINAIS	FASES INICIAIS	FASES FINAIS	FASES INICIAIS	FASES FINAIS
2ª fase - Ciências Contábeis	10ª fase - Engenharia de Materiais 9ª fase - ADM	2ª fase - Letras Libras 1ª fase - Odontologia	9ª fase - Odontologia	3ª fase - História 2ª fase - Relações Internacionais	9ª fase - Ciências Econômicas 8ª fase - Direito	3ª fase - Ciência e Tecnologia de Alimentos 2ª fase - Enfermagem 1ª fase - Design de Produto	9ª fase - Engenharia Eletrônica 7ª fase - Engenharia Química

Fonte: elaboração das autoras.

Sabendo que esta foi uma coleta piloto, percebemos que mesmo com essa disposição das células, não houve maiores prejuízos para a análise do fenômeno da concordância nominal. Dentre os procedimentos da coleta, os informantes preencheram uma ficha social com alguns dados pessoais, curso, fase do curso e localidade de residência. As entrevistas sociolinguísticas seguiram os moldes labovianos e foram conduzidas através de um roteiro previamente elaborado com tópicos sobre a família, o trabalho, a região onde mora, o curso, práticas de lazer etc. Alguns questionamentos gerais sobre o mercado de trabalho, política, segurança, educação também constavam no roteiro.

A busca pelos estudantes se deu pelos espaços do campus, onde abordávamos pessoas e convidávamos para colaborar em uma entrevista sobre a experiência da universidade. Não houve identificação como discentes de pós-graduação em Linguística a fim de evitar preocupações com normatização da fala e o monitoramento excessivo. É importante salientar que a busca por informantes e as entrevistas em si foram realizadas pelo grande grupo de pesquisadores da disciplina, de modo que a amostra foi compartilhada por todos para a realização de análises distintas acerca de diferentes fenômenos linguísticos.

As entrevistas foram realizadas, em sua maioria, na sala do projeto VARSUL (UFSC), no Centro de Comunicação e Expressão (CCE) no campus Florianópolis. Antes de iniciarmos as entrevistas, repassávamos com os informantes o Termo de Consentimento Livre e Esclarecido (TCLE), e pedíamos para que preenchessem os seus dados na ficha social. Os entrevistados tinham idades entre 17 e 30 anos. Em relação ao perfil socioeconômico, em geral, a maioria apenas estudava ou participava de projetos de iniciação científica como voluntários ou bolsistas.

Um fator que julgamos ser positivo para o trabalho foi a diversidade de cursos dos informantes encontrados para as entrevistas, essa heterogeneidade contribuiu positivamente

para a análise do perfil social dos acadêmicos da UFSC, ainda que fosse um roteiro único. Das 15 entrevistas realizadas, 14 são de cursos diversificados, sendo eles: Administração (deslocamento 1 - fases finais), Ciências Contábeis (deslocamento 1 - fases iniciais), Design de Produto (deslocamento 4 - fases iniciais), Direito (deslocamento 3 - fases finais), Ciências Econômicas (deslocamento 3 - fases finais), Enfermagem (deslocamento 4 - fases iniciais), Ciência e Tecnologia de Alimentos (deslocamento 4 - fases iniciais), Engenharia de Materiais (deslocamento 1 - fases finais), Engenharia Eletrônica (deslocamento 4 - fases finais), Engenharia Química (deslocamento 4 - fases finais), História (deslocamento 3 - fases iniciais), Letras Libras (deslocamento 2 - fases iniciais), Odontologia (deslocamento 2 - 1 de fases iniciais e 1 de fases finais) e Relações Internacionais (deslocamento 3 - fases iniciais).

O fenômeno linguístico investigado

O fenômeno selecionado para análise é a concordância de número, que aparece nas entrevistas de forma marcada e não marcada. Centrando o olhar na variação estilística, e dando ênfase apenas em condicionadores sociais/extralinguísticos, este estudo tem por objetivo demonstrar como a concordância nominal pode se transformar em um marcador estilístico estereotipado no meio acadêmico, o qual influencia integralmente a construção da persona universitária. Nesse sentido, o foco deste trabalho culmina na investigação do indivíduo, embora as entrevistas sociolinguísticas olhem para a comunidade. Esse tratamento foi escolhido por consideramos como relevantes os seguintes condicionadores sociais: a) o informante; b) o curso; c) o nível de letramento; e d) o deslocamento.

Algumas questões foram levantadas a fim de orientar a investigação, sendo: a) em que medida a marcação de concordância nominal expressiva no contexto acadêmico configura em um marcador estilístico da persona universitária?; b) em que medida há disparidade na realização da variável concordância nominal entre a fala de acadêmicos de início de curso em oposição à fala dos acadêmicos de final de curso?; c) de que maneira a marcação de concordância de número representa uma marca identitária representativa de estilo no meio acadêmico?; d) a fala monitorada, em termos de marcação de concordância de número, é majoritária em diferentes cursos de graduação ou restrita aos cursos mais elitizados?; e) como o perfil social dos informantes contribui para a realização de uma análise fundamentada na perspectiva estilística?

Com base nessas questões de investigação, algumas hipóteses foram levantadas, sendo elas: a) acadêmicos de fases iniciais e de fases finais de curso demonstram um alto índice de marcação de concordância nominal em suas falas; b) não há diferenças significativas entre a marcação de concordância e a não marcação na fala dos acadêmicos de início de curso e os acadêmicos de final de curso; c) a marcação expressiva de concordância nominal configura em um marcador estilístico constitutivo da persona universitária; e d) o perfil social do informante auxilia na delimitação da identidade social e linguística dos acadêmicos da UFSC.

Por fim, parte-se de uma tabulação quantitativa dos dados, objetivando identificar nas entrevistas realizadas as sentenças em que a concordância nominal foi marcada ou não marcada. Mediante essa contabilização, a análise dos dados se deu de forma qualitativa, justamente pela intenção de olhar para o estilo nesses usos. Nas páginas seguintes serão descritos e analisados alguns recortes dos 1.059 dados de concordância de número identificados na amostra. Os dados se dividem entre as variantes marcação (1.011 dados) e não marcação (48 dados) da variável concordância nominal¹⁴. A compilação dos dados se deu a partir da análise de fala das entrevistas realizadas.

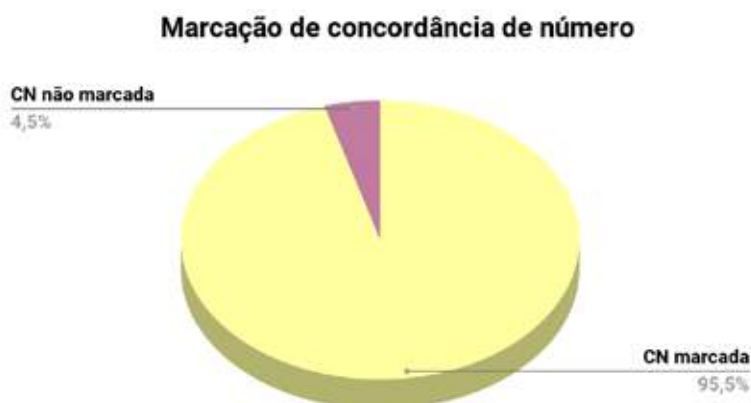
Análise qualitativa e estilística da marcação de concordância nominal

Salientamos que a nossa proposta de análise aborda qualitativamente os dados, discutindo, para além das marcas de variação mais recorrentes e típicas nos SN, formas variáveis relacionadas a algumas expressões idiomáticas e termos que apresentam variação na aplicação das regras de concordância nominal.

Um questionamento importante a ser feito é a função do estilo nesta proposta de discussão qualitativa dos dados. Se entendemos e assumimos o estilo como um fenômeno que marca a identidade linguística e social do falante, de que forma ele se relaciona com uma amostra extraída em uma comunidade de fala? Olhar qualitativamente para a variável concordância nominal através da perspectiva estilística significa entender as variantes marcação e não marcação como práticas identitárias dos informantes analisados. Em outras palavras, tratamos do estilo especialmente pelos resultados que os dados trouxeram: a expressiva marcação de concordância nominal.

Definidas as instâncias em que o estilo será explanado nesta análise, adentramos efetivamente nos dados extraídos da amostra. Em relação à variável concordância nominal, foram encontrados, como já mencionado, 1.059 dados, distribuídos em 1.011 dados com marcação e 48 dados sem marcação, conforme segue:

Gráfico 1 – ocorrência de concordância de número



Fonte: elaboração das autoras.

14 Ressaltamos que o foco da nossa análise não é quantitativo, trouxemos essas informações sobre o número de dados encontrados apenas para contabilizar as ocorrências identificadas na amostra.

De maneira geral, essas ocorrências de concordância marcada, as quais são a maioria, podem ser observadas nas seguintes construções:

- 1) “Eu sempre procurei morar perto da UFSC, nunca gostei de fazer *grandes deslocamentos diários*” (INFORMANTE 1);
- 2) “Você precisa que o país desenvolva *os seus próprios métodos sociais*” (INFORMANTE 1);
- 3) “[...] e Santiago é cercado por montanhas, então você vê *as montanhas cheias de neve*” (INFORMANTE 2);
- 4) “Tem *algumas páginas específicas* que eu gosto muito, tipo Quebrando o Tabu, 9GAG, Floripa Mil Grau, no Facebook” (INFORMANTE 3);
- 5) “É um dinheiro que dava para construir *várias escolas públicas*” (INFORMANTE 4);
- 6) “Eles leem muita coisa da internet, então é tudo muito abreviado, com *muitas gírias cibernéticas*” (INFORMANTE 5).

Nota-se que essas sentenças apresentam sintagmas nominais com a concordância marcada em todos os elementos, em todas as casas, isto é, nas posições pré-nucleares, nucleares e pós-nucleares. Reiteramos, então, que essas construções representam padrões de fala monitorada, oriundos de sujeitos que demonstram, em certo nível, consciência sobre a aplicação das regras de concordância de número.

No caso das ocorrências sem marcação de concordância, a minoria na amostra total, há sentenças em que a variação fica mais explícita e outras, em que é necessária a reflexão sobre as condições gramaticais de aplicação de regras de concordância. Olhando primeiro para as variações explícitas, os seguintes exemplos foram encontrados:

- 7) “Eu gosto de bicicleta. *A minhas atividade físicas* fica mais restritas a isso” (INFORMANTE 1);
- 8) “A minha mãe tem, o meu pai não. A minha mãe, *os livro* que ela consome [...]” (INFORMANTE 4);
- 9) “Eu acredito que fica meio que uma massificação, que vai gerar *consequências negativa*” (INFORMANTE 4);
- 10) “[...] vejo *as criança* da minha família, que eles têm uma dificuldade maior de dicção, uma dificuldade maior pra escrever” (INFORMANTE 5);
- 11) “[...] aqui a gente convive com pessoas de *vários estados diferente...* a gente muda um pouco o jeito de falar” (INFORMANTE 5);

12) “E a gente aprendia os xingamentos *das duas língua* pra se comunicar” (INFORMANTE 8);

13) “Não sei cozinhar... eu faço *uns miojo, umas pipoca, uns café* lá” (INFORMANTE 9);

14) “*As unidade básicas de saúde* têm algumas que já trabalham em um horário estendido” (INFORMANTE 10);

15) “*Pontos positivo*, que é um lugar que aceita todo mundo... não aceita todo mundo, é que são pessoas muito variadas” (INFORMANTE 11).

Tendo ciência de que em um SN o índice de marcação de concordância nominal¹⁵ é maior em elementos pré-nucleares e menor em elementos pós-nucleares (cf. SCHERRE, 1988, 1996 para estudos prévios sobre essa constatação), percebe-se nos trechos acima que algumas ausências de marcação despertam curiosidade ao passo que o informante marcou a concordância nos elementos pós-nucleares (cf. FERNANDES, 1996 para uma visão sobre esse tipo de dado). São exemplos desse tipo de ocorrência: *A minhas atividade físicas* (INFORMANTE 1) e *As unidade básicas de saúde* (INFORMANTE 10). As demais construções ficam restritas à marcação em elementos pré-nucleares, em sua maioria, e não marcação no núcleo do SN.

Em contrapartida a esses usos de não marcação mais explícitos, encontramos algumas sentenças em que a ausência de marcação pode ser explicada por outros fatores, senão os próprios gramaticais. Nas frases a seguir é possível refletir qualitativamente sobre os dados:

16) “*As pessoas* são mais *curta e grossas*” (INFORMANTE 1);

17) “Mas se você vai para o centro, por exemplo, você tem *bastante opções* para fazer” (INFORMANTE 4);

18) “Eu tenho *bastante lugares* que eu gostaria de conhecer, tanto na América quanto na Europa” (INFORMANTE 4);

19) “Ela lê *bastante livros* de padre” (INFORMANTE 4);

20) “Eu gosto muito de comer lasanha, *bastante tipos*, presunto e queijo principalmente” (INFORMANTE 4);

21) “No curso de licenciatura como eu tô, a gente tem *bastante disciplinas* para isso” (INFORMANTE 4);

15 Estudos como o de Guy (1981) e Brandão e Vieira (2012) postulam que a 1ª posição no SN é o lugar privilegiado para marcação de concordância de número no português brasileiro. Em dados sobre o português paulistano Oushiro (2015, p. 402) também confirma esse aspecto: “[...] verifica-se uma distribuição extrema quando a palavra ocupa a primeira posição: a taxa de marca zero nesse contexto é de apenas 0,3%”.

- 22) “Ai eu acho assim que o CCB tá caindo *aos pedaço*, tá caindo *aos pedaço* assim, tá ridículo” (INFORMANTE 12);
- 23) “Quem vem direto, assim, da rede pública, sem ter passado por nenhum ensino técnico nem nada, essas pessoas elas têm *bastante dificuldades* no começo da faculdade (INFORMANTE 8);
- 24) “Eu só estudando já faço um negócio meio *nas coxa*, imagina se eu tivesse que trabalhar” (INFORMANTE 9);
- 25) “Na odonto a gente tem *bastante colegas negros* e diminui sim” (INFORMANTE 10).

Diante desse cenário, constata-se que as formas variantes *curta e grossas, aos pedaço e nas coxa* representam expressões idiomáticas, as quais podem ter sido aprendidas e armazenadas pelos informantes justamente nesta configuração e não com a concordância marcada, o que fundamentaria a não utilização das regras de concordância de número¹⁶. Já as construções que envolvem o uso de *bastante* adentram em outro debate. Especificando, nas gramáticas normativas tradicionais (CUNHA; CINTRA, 2013), está posto que quando o termo *bastante* precede substantivos no plural e exerce a função de adjetivo ou de pronome indefinido, deve sofrer variação, ficando invariável somente quando corresponder ao uso de um advérbio de intensidade e funcionar como o substituto de *muito*. No entanto, sabendo dos usos que são feitos desse marcador no português brasileiro, é comum que os falantes tenham internalizado apenas o uso do *bastante* sem a marca de plural, o que justificaria a sua ocorrência na fala de mais de um informante, embora fosse o caso de se utilizar *bastante* com variação.

Oushiro (2015) discorre sobre o uso da expressão “dois pastel e um chopes” ser um estereótipo da fala paulistana, especialmente a fala da região do bairro do Mooca. A autora aponta que nesta expressão há marca zero de plural em “pastel” e um -s adicional em “chopes”, tornando possível a discussão da hipótese de que os moradores do bairro Mooca tendem à marca zero em relação à concordância nominal (p. 393). Nas entrevistas analisadas por Oushiro (2015, p. 396), o uso de “dois pastel e um chopes” gerou comentários metalinguísticos sobre a marca zero de CN ser sinônimo de “não saber falar português”. A autora pontua ainda que as avaliações que associam o uso dessa expressão ao falar paulistano, à imigração italiana e a certos bairros parecem se restringir aos falantes mais velhos. Já os falantes mais jovens tendem a associar o uso dessa expressão com baixa escolaridade (OUSHIRO, 2015, p. 400). Segundo a

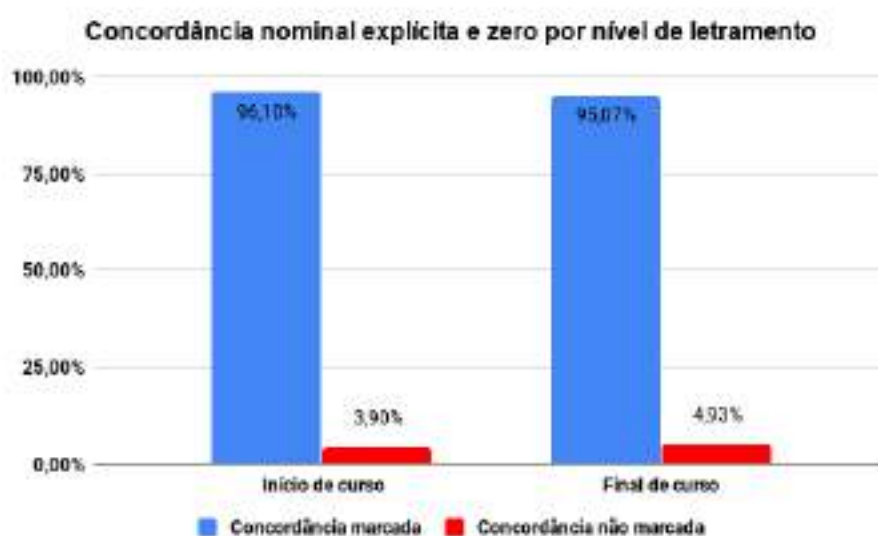
16 Agradecemos a sugestão do parecerista anônimo para explorarmos outras hipóteses relacionadas ao uso dessas expressões idiomáticas, como o lugar em que elas ocupam na língua em relação ao grau de monitoramento. Reiteramos que a investigação desse aspecto é uma motivação de análise futura, que foge do escopo deste trabalho. Fulgêncio (2008) apresenta uma abordagem sobre expressões fixas e idiomatismos no PB, trazendo uma discussão acerca das expressões idiomáticas representarem um tipo de expressão fixa no léxico.

autora, alguns informantes também assumem que a expressão “dois pastel e um chopes” é uma marca de identidade paulistana. No entanto, Oushiro (2015) comenta que a fala dos mais jovens parece indicar que a associação estereotipada da expressão “dois pastel e um chopes” com os paulistanos pode estar em vias de desaparecimento (p. 401).

Para Oushiro (2015), os comentários metalinguísticos analisados acerca dessa expressão, bem como os padrões de variação na comunidade investigada revelam que certos usos linguísticos podem ser associados a certas identidades sociais: “O estudo do significado social da variação, para além dos amplos padrões observados, pode conduzir a novas interpretações da vitalidade de certas formas linguísticas em diferentes comunidades” (p. 422). Apesar de não encontrarmos esse cenário em nossos dados, tendo em vista que Oushiro (2015) trabalhou com a variação na CN e com a associação da marca zero a um grupo específico, discutiremos a seguir que a alta marcação de concordância nominal na comunidade acadêmica investigada também pode funcionar como objeto de análise do significado social da variação. Nesse sentido, a expressiva marcação de concordância pode contribuir para a criação de um estereótipo da persona universitária, passando a existir uma associação entre um grupo (universitários) e um uso linguístico (marcação expressiva de CN). Há uma projeção de expectativa, portanto, de que as marcas zero de CN sejam mais frequentes na fala de sujeitos que não estão inseridos no meio acadêmico, o que reforça o estereótipo da persona universitária.

Apresentadas essas questões gerais sobre os dados encontrados, partimos agora para uma tentativa de responder a uma de nossas questões de investigação: em que medida há disparidade na realização da variável concordância nominal entre a fala de acadêmicos de início de curso em oposição à fala dos acadêmicos de final de curso? Foram encontradas 394 observações de concordância nominal marcada na fala de informantes de início de curso e 16 dados de não marcação; em oposição a 617 ocorrências de marcação em final de curso e 32 dados de não marcação. Esses números devem ser tomados em proporcionalidade:

Gráfico 2 – realização da concordância de número por nível de letramento

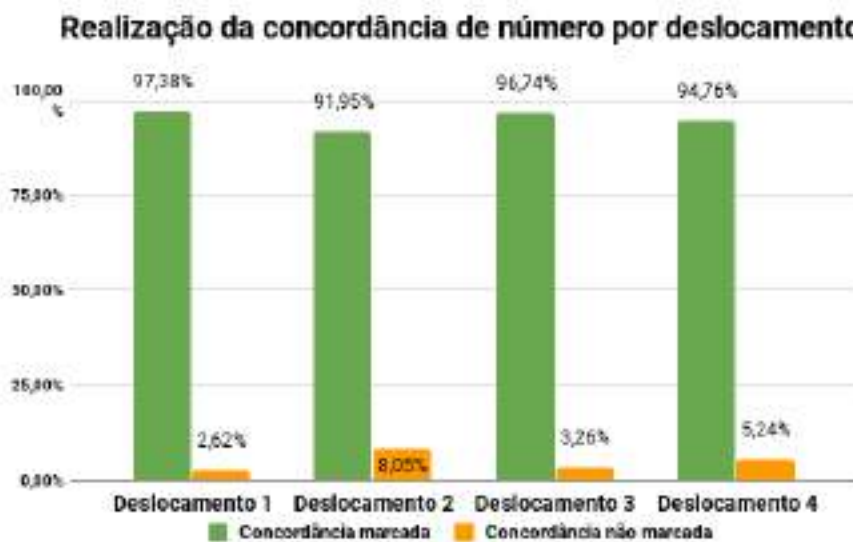


Fonte: elaboração das autoras.

Relativamente, a proporção entre os dados de início e final de curso é semelhante, entretanto, deve-se ter cautela ao comparar de maneira direta as ocorrências nos dois níveis de letramento, em função da quantidade de dados não ser a mesma, com uma diferença, conforme citado, de quase o dobro entre um e outro. Isto significa que não há uma disparidade grande entre os dados de concordância nominal dos informantes de início de curso e os dados de concordância dos informantes de final de curso, ambos marcam e apresentam ausência de marcação de forma semelhante. Seguindo essa lógica, cabe ser enfatizado que além da expressiva marcação de concordância na amostra, no conjunto de observações analisadas também é possível reconhecer que a fala dos informantes indica o seu perfil de sujeitos marcadores de CN, independentemente do nível de letramento em que estão inseridos. Oushiro (2015, p. 421) evidencia que: “[...] a associação de traços linguísticos a grupos de indivíduos reside em suas *frequências* de uso, e não em *tendências*”. Sobre esse ponto, salientamos que os acadêmicos da UFSC que compuseram a amostra são agentes que monitoram a sua fala, o que acarreta pouca não marcação de CN, isto para os dois níveis de letramento: início e final de curso, comprovando uma das hipóteses levantadas. Assim, podemos destacar que esses acadêmicos usam as variantes marcação e não marcação de CN na mesma proporcionalidade.

Quanto aos deslocamentos, apresentaremos no Gráfico 3 um movimento comparativo entre as ocorrências de marcação e não marcação por célula social. Ainda que os deslocamentos não tenham sido preenchidos na mesma proporção (ver Quadro 1), por ser uma coleta piloto, consideramos pertinente ilustrar de que maneira os informantes expressam CN em cada célula social, conforme segue:

Gráfico 3 – concordância de número expressa por célula social (deslocamento)



Fonte: elaboração das autoras.

Novamente, a noção de proporcionalidade se torna importante para entender como os dados estão sendo postos nesta análise. Em questão de números, a expressão da concordância pode ser esquematizada da seguinte forma: Deslocamento 1: 186 observações de marcação

e 5 de não marcação; Deslocamento 2: 137 observações de marcação e 12 de não marcação; Deslocamento 3: 326 observações de marcação e 11 de não marcação; Deslocamento 4: 362 observações de marcação e 20 de não marcação. No Gráfico 3, o deslocamento 1 apresenta maior diferença em questão de pontos percentuais entre marcação e não marcação, por refletir o resultado proporcional aos dados dispostos acima. Já no deslocamento 2¹⁷ é onde se encontra a maior frequência de não marcação de concordância nominal.

Outro ponto relevante a ser explorado é a alternância de monitoramento ao longo da entrevista. No início, os informantes responderam às perguntas seguindo um protocolo mais formal, o que impulsionou altos índices de marcação da concordância nominal. Ao passo que a entrevista foi se estendendo e o informante se acostumou com a situação comunicativa da qual estava participando, o monitoramento começou a decair e algumas formas variáveis vieram à tona, dando lugar às construções com maiores ocorrências de não marcação. As temáticas também foram importantes para se pensar essa alternância de monitoramento, uma vez que quando as perguntas enfocavam a prática acadêmica, o curso do universitário em si, poucas observações de não marcação foram constatadas. Em oposição, quando os questionamentos adentravam em assuntos como lazer, família, culinária, *hobbies* e lembranças, o nível de monitoramento passava a ser mais baixo, dando margem para construções como:

26) “Em outros lugares sempre tem *uns cara* que passam muito do limite, mano”
(INFORMANTE 12);

27) “Faz uns três anos eu acho... É que eu fui meio retardado eu acho... sai correndo atrás dele, sei lá... tu tá na tua casa e tem um cara tentando roubar *tuas coisa*” (INFORMANTE 7);

28) “As pessoas mudam também né, a vida muda também, então acho que mudou bastante... e memórias boas eu tenho muitas assim: comendo coxinha *cus muleque* tudo sentado no meio fio, então acho que essas memórias são as que mais ficam” (INFORMANTE 9);

29) “Essa é a diferença, então lá sempre tem alguma coisa aberta, *os lanche* são mais baratos... cê comprava um copo de coxinha no ponto de ônibus por R\$1,00” (INFORMANTE 9);

30) “E aí geralmente reúne toda a família assim, pra descascar milho, pra catar pimenta no pé, é *uns programa* de índio assim” (INFORMANTE 9).

17 Um aspecto relevante sobre os deslocamentos é o fato de que a maior frequência de não marcação é atestada no deslocamento 2, em que se localizam informantes que cursam Odontologia, um dos cursos tidos como elitizados na universidade. Há a projeção de uma expectativa de maior monitoramento na fala desses informantes e o uso de uma variedade com pouca variação na aplicação das regras de concordância nominal, conforme será discutido adiante.

Ainda neste contexto, a informante 9 despertou curiosidade ao totalizar 13 ocorrências de não marcação. Faz-se necessário mencionar que essa informante estabelecia certo grau de intimidade com a entrevistadora, já que se conheciam de outras vivências. Mediante esse fator, a queda no monitoramento pode ser justificada à medida que foi possível reconhecer que ela se sentia à vontade durante a entrevista. Em vários momentos houve descontração entre ambas e a partir disso é que as observações de não marcação emergiram. Pode-se dizer que o vernáculo foi alcançado, sobretudo, quando voltamos a analisar a seguinte construção:

31)“Não sei cozinhar...eu faço *uns miojo, umas pipoca, uns café lá*”
(INFORMANTE 9).

Neste instante, entrevistadora e entrevistada dialogavam sobre o tempo em que moravam juntas em uma república universitária, lembrando as receitas culinárias que fizeram na companhia uma da outra. Ademais, além do grau de intimidade, gostaríamos de destacar que o fato dessa informante apresentar mais marcas de não marcação de CN pode ser fruto de suas próprias práticas identitárias de uso da língua, revelando a sua prática estilística, que além de marcas zero de concordância de número, também é acompanhada de gírias e expressões idiomáticas, demonstrando, por ora, traços locais, estilísticos e identitários.

A concordância nominal como um marcador estilístico

Sabendo que os acadêmicos-informantes estão inseridos no contexto universitário, acaba sendo previsível que eles marquem mais a concordância nominal do que outros sujeitos que não tiveram acesso ao mesmo espaço de letramento. A concordância nominal entra nesta análise como um marcador estilístico à medida que se relaciona também com o perfil social do informante-falante. Falamos da posição de acadêmico e das significações que esses universitários trazem consigo ao serem alunos de dado curso e estarem em dada fase/nível de letramento. Isto posto, destacamos que a principal justificativa para relacionar o estilo e a concordância nominal está na elitização dos cursos que a maioria dos informantes frequenta. Junto de um curso de elite, possivelmente, está um acadêmico que teve acesso às práticas e eventos de letramento que contribuíram para o armazenamento das regras de marcação de concordância nominal, o que reflete em seus usos linguísticos na universidade.

O estilo aqui é reflexo da posição social de universitário e de acadêmico de cursos, por vezes, elitizados. Recuperando o que já fora dito, quando Eckert (2005, 2012) afirmou que estilos emergem de posturas repetidas, referia-se tanto ao linguístico quanto ao social. Partindo dessa premissa, ao olhar para o perfil social do informante através de seus próprios discursos e outras informações, é possível identificar que um estudante de um curso elitizado pode ter tido acesso a uma rotina com práticas de letramento que possibilitam um fazer linguístico mais monitorado, independente do contexto em que esteja circulando. Não é só a universidade que

Diadorim, Rio de Janeiro, vol. 24, número 1, p. 203 - 232, 2022.

aguça a marcação de concordância, mas as próprias práticas com a língua que estes sujeitos exerceram anteriormente ao longo de sua história, em diferentes situações comunicativas.

Camacho e Salomão-Conchalo (2016) discutem a relação entre identidade e a marcação variável de CN em comunidades de prática de uma escola pública. Os autores apontam que essa escola é caracterizada por diferentes comunidades de prática, sendo as mais representativas os *funkeiros* e os *ecléticos*. De acordo com os autores, ambos os grupos demonstram diferenças em seus usos linguísticos, o que pode ser moldado pela própria forma de se relacionar com as outras comunidades de prática e com a instituição em que estão inseridos: a escola. Os autores se interessam, sobretudo, pela análise das dinâmicas e das práticas sociais desses dois grupos, que podem refletir o uso variável da CN. Para Camacho e Salomão-Conchalo (2016), a variação de pluralidade no SN pode constituir um campo de significados sociais, o qual pode incidir sobre a construção de identidade. Em relação à variável CN, os autores pontuam que: “[...] o processo variável de concordância nominal nas comunidades de prática faz parte de uma produção estilística ativa de diferenciação social” (CAMACHO; SALOMÃO-CONCHALO, 2016, p. 46). Quanto aos dados quantitativos discutidos para as comunidades de prática dos *ecléticos* e dos *funkeiros*, a aplicação da regra de CN ocorreu em 92,3% dos casos possíveis para os *ecléticos* e em 62% dos casos para os *funkeiros* (CAMACHO; SALOMÃO-CONCHALO, 2016, p. 55). Segundo os autores, a constituição das comunidades de prática e o perfil social dos membros auxiliam na interpretação do significado social da aplicação variável de CN. Para os membros *funkeiros*: “A variação de pluralidade é apenas um recurso estilístico de construção de identidade, ao indexar distanciamento de outros grupos, e mesmo um perfil individual e próprio, que se destaca dos demais membros do mesmo grupo” (CAMACHO; SALOMÃO-CONCHALO, 2016, p. 60). Já para os membros *ecléticos* “[...] aplicar positivamente a regra de CN significa emblematicamente aproximar-se do que é aceitável pela norma padrão, além de assegurar o distanciamento seguro dos funkeiros em relação às diferenças de posição ideológica” (CAMACHO; SALOMÃO-CONCHALO, 2016, p. 60). Na visão dos autores, trata-se de um aproveitamento estilístico que atua na indexação de identidades. Em nossos dados a expressiva marcação de concordância nominal foi interpretada como constitutiva do estilo dos universitários e como um aspecto identitário dessa persona.

Para demonstrar o impacto das práticas sociais na construção da persona universitária, começamos pela própria valorização que os cursos de graduação que os acadêmicos frequentam possuem na sociedade. O Quadro 2 apresenta informações sobre os cursos dos informantes em relação às notas de corte do SISU e o índice de concorrência no vestibular da UFSC:

Quadro 2 – notas de corte do SISU-UFSC e índice de candidato por vaga no vestibular da UFSC

NOTA DE CORTE SISU-UFSC E ÍNDICE DE CANDIDATOS POR VAGA		
CURSO	NOTA DE CORTE SISU	CANDIDATOS x VAGA
ADMINISTRAÇÃO (DIURNO)	746.21	9,77
CIÊNCIAS CONTÁBEIS (NOTURNO)	752.53	7,42
DESIGN DE PRODUTO	727.83	6,5
DIREITO (DIURNO)	817.68	35.42
CIÊNCIAS ECONÔMICAS (NOTURNO)	761,54	6,74
ENFERMAGEM	779.67	22,04
CIÊNCIA E TECNOLOGIA DE ALIMENTOS	696.71	1,96
ENGENHARIA DE MATERIAIS	772.07	3,04
ENGENHARIA ELETRÔNICA	781.22	5,29
ENGENHARIA QUÍMICA	788.43	22,82
HISTÓRIA (NOTURNO)	747.15	8,63
LETRAS LIBRAS	672.69	4,14
ODONTOLOGIA	771.31	23,83
RELAÇÕES INTERNACIONAIS	785.3	19,14

Fonte: elaboração das autoras com base nas informações divulgadas pela Coperve em 2019 e 2020.

No que concerne à elitização dos cursos de graduação frequentados pelos informantes, observa-se que nem todos apresentam um índice de concorrência de candidato por vaga alto, mas todos passam por uma nota de corte relativamente alta se o ingresso na universidade for pelo SISU. Isto revela que a própria UFSC, enquanto universidade pública federal, é uma instituição de difícil acesso, necessitando, em muitos casos, até mesmo comentados pelos próprios informantes, de uma preparação em cursinhos de pré-vestibular. Se a universidade é de difícil acesso e os próprios universitários reconhecem isso, é porque estamos falando de um meio acadêmico que traz em seu interior cursos prestigiados, que se tornam elitizados na sociedade, como pode ser constatado, sobretudo, pelos dados dos cursos de: Direito, Enfermagem, Engenharia Química, Odontologia e Relações Internacionais, cuja concorrência é significativa tanto no SISU quanto no vestibular da UFSC.

Os cursos dos quais os acadêmicos advêm são um dos fatores que devem ser considerados acerca do seu perfil social. Torna-se necessário salientar também que independente de alguns cursos serem mais prestigiados e elitizados do que os outros, conforme apontado acima, a marcação de concordância continua expressiva na fala dos entrevistados dos outros cursos de

graduação, de modo que essa ocorrência não se restringe somente aos cursos mais valorizados. Isto significa que a fala monitorada, em termos de aplicação das regras de concordância de número, é majoritária entre os entrevistados. A marcação da concordância nominal é um elemento já enraizado na cultura acadêmica, mediante a função social de sujeito letrado acarretada ao acadêmico ao se inserir na universidade.

Ao discutir a relação entre o conceito de identidade e pesquisas sociolinguísticas, Oushiro (2019) aponta que as identidades se relacionam com o social: “As diferentes definições de identidade ressaltam o caráter relacional do conceito, que deve passar, necessariamente, pela relação entre um indivíduo e construtos sociais mais amplos” (OUSHIRO, 2019, p. 322). Logo, pode-se refletir que a identidade dos usos linguísticos pode ser atravessada pelo papel social que os falantes exercem no meio em que circulam. Nesse sentido, a expressiva marcação de concordância na fala dos acadêmicos da amostra demonstra que os seus usos linguísticos, suas práticas estilísticas, são antes atravessados pelo papel social de acadêmico. Ao propormos a discussão sobre a criação de uma persona universitária, adentramos em aspectos identitários e usos linguísticos que caracterizam essa persona.

Assim sendo, os universitários são vistos como indivíduos que controlam e monitoram a sua fala, o que leva à expressiva marcação de concordância. Esta associação entre o perfil de acadêmico e o uso de uma fala monitorada, com usos linguísticos mais próximos ao padrão, é um estereótipo que se consolidou na sociedade, já que esses indivíduos têm acesso às práticas e eventos de letramento diferenciadas, distintas das dos sujeitos que não frequentam a universidade. Nesse sentido, ocupar a posição de estudante universitário implicaria a aplicação das regras de concordância durante a língua em uso. Portanto, a variável concordância de número entra neste jogo de linguagem como um marcador estilístico que constitui a persona universitária e pode ser explicado por três vertentes: a) ao se revelar expressiva na própria fala dos estudantes, sendo uma postura linguística repetida de forma contínua; b) ao estar atrelada com a sua função social de acadêmico; c) ao alimentar o estereótipo criado para diferenciar aquele que está dentro da universidade e aquele que está à margem, fazendo com que o primeiro aplique as regras gramaticais de concordância nominal, sendo esperado que ele as utilize pela sua posição na hierarquia social, e o segundo esteja à margem da academia, sendo esperado que ele realize mais formas variáveis.

A elitização de alguns cursos, o difícil acesso às universidades federais e até mesmo o estereótipo do que é ser um acadêmico contribuem para a manutenção da ideia de que um universitário deve fazer uso de uma linguagem mais próxima à norma culta. Neste entremeio, incorpora-se desde a construção de frases coesas até a utilização de regras de concordância, seja ela nominal ou verbal. Os próprios informantes demonstram consciência da propagação da normatização no meio social em que circulam:

*Na escola até, pelo menos na educação que eu tive, teve um cuidado de dizer: ah, essa é a norma culta, mas não significa que as variantes linguísticas são ruins. Socialmente a gente sabe que se o cara falar 'praneta' as pessoas olham com preconceito ou olham feio né. **Isso ficou até bem mais forte no direito**, porque é incrível como o pessoal gosta de analisar o estilo de escrita da pessoa. Se a pessoa escreve um erro da norma culta bem forte no português as pessoas já julgam... já pensam ó, esse cara não sabe o que tá falando, esse cara é um ignorante... **Tem essa mentalidade muito forte** (INFORMANTE 6, grifos nossos).*

*Porque eu sempre aprendi que existe a forma correta gramatical, mas **a parte linguística não existe a forma correta de falar né...** tipo, tu fala de acordo com o lugar que tu nasceu, que tu vive (INFORMANTE 7, grifos nossos).*

*Incomoda? Incomoda! Mas é porque **alguém te corrigiu algum dia pra incomodar**, te incomoda porque **alguém te ensinou que aquilo ia incomodar e tá errado**, por isso que te incomoda. Mas fora isso, **nada fora da normalidade... é um incomoda que tipo... social mesmo** (INFORMANTE 9, grifos nossos).*

Tendo consciência de que a língua varia, independente de existir uma norma padrão, os informantes trazem em sua fala a percepção de que até mesmo os acadêmicos podem fazer uso de formas variáveis, indicando a apropriação de variedades com diferentes valores sociais. Outro ponto que merece atenção é o fato de os três informantes trazerem à tona a ideia de julgamento social ao se fazer uso de formas em variação. Estes indivíduos reconhecem que a língua não é homogênea, sendo constituída de práticas identitárias e heterogêneas, ligadas às características dos próprios sujeitos. Ao mencionarem que as pessoas utilizam variedades locais, relacionadas às suas regiões de origem; e ao salientarem que a sociedade julga e se incomoda com quem faz uso de formas em variação, por saberem da exigência de se utilizar a norma culta, estão, automaticamente, evocando as concepções de adequação, identidade e práticas locais, componentes da perspectiva estilística.

Ao refletirmos sobre a citação do acadêmico de direito (informante 6), pontuamos duas reflexões importantes para este estudo: 1) efetivamente existe um estereótipo quanto ao estilo dos acadêmicos, de modo que se espera que eles façam o uso da língua mais próximo da norma culta possível, expresso especialmente na construção: “Se a pessoa escreve um erro da norma culta bem forte no português as pessoas já julgam... já pensam ó, esse cara não sabe o que tá falando, esse cara é um ignorante”. Este estereótipo é reforçado tanto pelos sujeitos que não estão inseridos na academia, quanto pelos acadêmicos ao se sentirem, ainda que não intencionalmente, pressionados a fazerem uso de uma linguagem monitorada, que segue o padrão. No que tange à noção de “erro” ao qual o informante se refere, podemos trazer a própria concordância de número para o debate, tendo em vista que a não marcação da concordância nominal é um fenômeno que pode gerar estigma no meio social.

O segundo ponto de reflexão é: 2) o perfil social é importante para se traçar as características estilísticas de uma persona, de modo que este informante que advém de um curso elitizado - o

direito - apresenta consciência de que ser um acadêmico de direito implicaria não fazer uso recorrente de formas em variação. Mais uma vez entramos na questão do estereótipo, só que por outra ótica: aqui o acadêmico vem de um curso prestigiado, em que se idealiza uma persona universitária letrada. A relação entre o valor atribuído a um curso de graduação, o estilo e a construção de uma persona universitária se traduzem nos seguintes trechos: “Isso ficou até bem mais forte no direito. [...] Tem essa mentalidade muito forte”. Por consequência, refletimos que ser um acadêmico imerso na cultura do direito, ser um acadêmico de um curso, por vezes, tido como elitizado, pode acarretar a expectativa de que se ocupe a posição de falante ideal, que utiliza uma língua invariável e homogênea. Isto implicaria preservar continuamente um fazer linguístico sem o uso de variantes com valor social desprestigiado. No que diz respeito à variável deste trabalho, implicaria fazer uso de uma fala marcada por concordâncias de número realizadas, independente da situação comunicativa.

Em uma última instância, evidenciamos que os informantes têm consciência das escolhas linguísticas que fazem ao usarem a língua, e das escolhas que esperam que eles façam mediante a posição de universitários - sujeitos altamente letrados - que ocupam na sociedade. Trabalhar com perfil social a partir dos cursos de graduação possibilitou que abordássemos o condicionador informante, uma vez que traçamos fatores importantes para que se compreenda o movimento de construção da persona universitária através das atribuições sociais que os acadêmicos carregam. A inserção no meio acadêmico fomenta a criação de um estereótipo do que é ser universitário e quais funções estes sujeitos sociais devem exercer com a língua. Se é esperado que os acadêmicos, por estarem imersos na academia, utilizem uma variedade mais próxima da norma culta, também é esperado que eles apliquem as regras de concordância de número, marcando integralmente os elementos dos SN durante os seus registros orais e escritos. Não só é esperado, como foi confirmada a expressiva marcação de CN na fala dos informantes-acadêmicos, fazendo com que o uso de formas em variação observado ao longo das entrevistas fosse pouco expressivo em um cenário com tantas marcações.

As discussões propostas ao longo desta seção se alinham com as hipóteses previamente estabelecidas neste estudo: a concordância nominal é um marcador estilístico de construção da persona estereotipada do acadêmico da UFSC à medida que o estilo desses sujeitos está estritamente ligado à função social de acadêmico, em que se prevê um uso culto da língua. A expressiva marcação de concordância nominal pode ser interpretada como uma postura linguística repetida, relacionada à cultura social em que os acadêmicos circulam e atuam. Cabe indicar que o processo de construção de uma persona não depende exclusivamente das escolhas estilísticas que o falante faz, mas também da visão exterior do outro, que avalia e atribui significação às práticas sociais e linguísticas exercidas.

Considerações finais

As pesquisas realizadas no campo da Sociolinguística Variacionista, com foco na descrição e análise linguística, têm sido um importante material de reflexão sobre a língua em uso e a sua associação com os contextos sociais. Quando assumimos a língua como objeto de estudo, entendendo a sua estrutura heterogênea e sistemática, e considerando-a parte do contexto social dos falantes, podemos perceber que são muitos os fatores que atuam sobre ela, condicionando sua dinamicidade. Por conta disso, a reflexão e análise realizada nesta pesquisa se apoiam nos estudos de terceira onda, já que esse ciclo volta o olhar para o falante. Em nosso estudo investigamos a fala de acadêmicos de alguns cursos da UFSC, buscando entender as práticas linguísticas que constituem a sua persona, adentrando na questão do estilo.

Partindo do viés da terceira onda, percebemos que a construção da persona universitária (dos acadêmicos da UFSC) tem como forte característica a expressiva marcação de concordância nominal enquanto prática identitária. Ainda que a atenção voltada aos dados tenha sido qualitativa, não podemos deixar de destacar que dos 1.059 dados encontrados, 95,5% apresentaram marcação de concordância de número. Obter esse resultado em um estudo que contou com universitários de cursos distintos, com diferentes níveis de letramento e oriundos de deslocamentos diversos é bastante significativo para as reflexões sobre construção da persona universitária e seus usos linguísticos.

Quando apresentamos as taxas de proporcionalidade, tanto em nível de letramento (início e final de curso) como de deslocamento, confirmamos que a marcação da concordância nominal é um fator indicativo de construção de um perfil acadêmico de monitoramento da fala. Do mesmo modo, consideramos a elitização de alguns cursos um fator social importante na relação entre estilo e concordância nominal. Explica-se à medida que acadêmicos de cursos concorridos e com altas notas de corte no SISU, provavelmente, já entraram na universidade com um alto nível de letramento, em decorrência de práticas sociais e linguísticas anteriores. A construção de uma persona universitária, associada a certos usos linguísticos, pode ser reflexo de um estereótipo cristalizado na sociedade, o qual espera comportamentos linguísticos específicos dos sujeitos que frequentam a universidade.

Tendo como base os resultados da análise e todos os aspectos levantados, afirmamos que a marcação de concordância nominal pode ser interpretada como um marcador estilístico, cuja expressão é fundante no processo de construção identitária da persona universitária. Estilos emergem de posturas linguísticas e sociais, pessoas adotam estilos, fazendo e refazendo as suas identidades de maneira contínua.

Referências

BRANDÃO, S. F. Concordância nominal em duas variedades do português: convergências e divergências. *Revista Veredas*, v.15, n.1, p.164-178, 2011.

BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. *Alfa*, v. 56, n.3, p. 1035-1064, 2012.

CAMACHO, R. G. Uma reflexão crítica sobre a teoria sociolinguística. *Revista Delta*, v.26, n.1, p.141-162, 2010.

CAMACHO, R. G.; SALOMÃO-CONCHALO, M. H. A variação de plural no SN como um indexador de identidade. *Todas as Letras*, v.18, n.2, p.46-63, 2016.

CARVALHO, H. M. de. *Concordância nominal: uma análise variacionista*. João Pessoa: UFPB, 1997. Dissertação (Mestrado em Letras), UFPB, João Pessoa, 1997.

COPERVE. “Relação Candidatos/Vaga”. *Vestibular UFSC 2019*, Florianópolis. Disponível em: <<https://vestibular2019.ufsc.br/relacao-candidatos-vaga/>>. Acesso em 15 de fevereiro de 2020.

COPERVE. Notas Máximas e Mínimas – Classificados: UFSC/SISU 2020. *Seleção UFSC/SISU 2020*. Disponível em: <<https://sisu2020.ufsc.br/>>. Acesso em 15 de fevereiro de 2020.

CUNHA, C; CINTRA, L. *Gramática do português contemporâneo*. 6ª ed. Rio de Janeiro: Lexikon, 2013.

ECKERT, P. *Linguistic variation as social practice*. Oxford: Blackwell, 2000.

_____. *Variation, convention and social meaning*. Paper Presented at the Annual Meeting of the Linguistic Society of America. Oakland, 2005.

_____. *Three waves of variation study: the emergence of meaning in the study of sociolinguistic variation*. *Annual Review of Anthropology*, Palo Alto, n.41, p.87-100, 2012.

ECKERT, P.; MCCONNELL-GINET, S. Comunidades de práticas: lugar onde coabitam linguagem, gênero e poder (1992). In: OSTERMANN, A. C.; FONTANA, B. (org.). *Linguagem, gênero, sexualidade: clássicos traduzidos*. São Paulo: Parábola, 2010, p. 93-108.

FERNANDES, M. *Concordância nominal na Região Sul*. Dissertação (Mestrado em Linguística) do Programa de Pós-Graduação em Linguística, UFSC, Florianópolis, 1996.

FONSECA; FRANCESCHINI; LOREGIAN-PENKAL. Análise da concordância nominal em Guarapuava, Paraná. *Revista Interfaces*, v. 9, n. 2, p. 127-140, 2018.

FREITAG, R. M. K.; MARTINS, M. A.; TAVARES, M. A. Bancos de dados sociolinguísticos do Português Brasileiro e os estudos de terceira onda: potencialidades e limitações. *Revista Alfa*, v.56, n.3, p. 907-934, 2012.

FREITAG, R. M. K. Banco de dados Falares Sergipanos. *Working Papers em Linguística*, v.14, n. 2, p. 156-164, 2013.

FULGÊNCIO, L. *Expressões fixas e idiomatismos do português brasileiro*. Belo Horizonte: PUC Minas, 2008. Tese (Doutorado em Linguística e Língua Portuguesa) do Programa de Pós-Graduação em Letras, PUC Minas, Belo Horizonte, 2008.

GUY, G. R. *Linguistic variation in Brazilian Portuguese: aspects of the phonology, syntax and language history*. 1981. Thesis (PhD), University of Pennsylvania, Philadelphia, 1981.

GUY, G. R.; ZILLES, A. *Sociolinguística Quantitativa: instrumental de análise*. São Paulo: Parábola Editorial, 2007.

GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (org.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise*. Coleção Linguística. v.3. Florianópolis: Insular, 2014.

GÖRSKI, E. M.; VALLE, C. R. M. A variação estilística em entrevistas sociolinguísticas: uma (re)leitura do modelo laboviano. In: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (org.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise*. Coleção Linguística. v.3. Florianópolis: Insular, 2014, p. 67-92.

HORA, D. da. Estilo: uma perspectiva variacionista. In: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (org.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise*. Coleção Linguística. v.3. Florianópolis: Insular, 2014, p. 19-30.

LABOV, W. The social motivation of a sound change. *Word*, v.18, p.1-42, 1963.

_____. *The Social Stratification of English in New York City*. Washington, DC: Cent. Appl. Ling, 1966.

_____. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972a.

_____. *Language in the inner city: studies in Black English Vernacular*. Philadelphia: University of Pennsylvania Press, 1972b.

_____. *Principles of linguistic change*. vol. 3: Cultural and cognitive factors. Oxford: Wiley-Blackwell, 2010.

LUCCHESI, D.; DÁLIA, J. Equacionando o efeito da posição na variação da concordância nominal de número. *Revista Linguística*, v.16, número especial comemorativo, p. 771-798, 2020.

MILROY, L. *Language and Social Networks*. Oxford: Blackwell, 1980.

OUSHIRO, L. Dois pastel e um chopes: a concordância nominal e identidade(s) paulistana(s). *Revista de Estudos da Linguagem*, v. 23, n. 2, p. 389-424, 2015.

_____. Conceitos de identidade e métodos para seu estudo na sociolinguística. *Estudos Linguísticos e Literários*, n. 63, número especial, p. 304-325, 2019.

SALOMÃO-CONCHALO, M. H. *A variação estilística na concordância nominal e verbal como construção de identidade social*. São José do Rio Preto, UNESP, 2015. Tese (Doutorado em Estudos Linguísticos) do Programa de Pós-Graduação em Estudos Linguísticos, UNESP, São José do Rio Preto, 2015.

SEVERO, C. G. Estilo, variação linguística e discurso. In: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (org.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise*. Coleção Linguística. v.3. Florianópolis: Insular, 2014, p. 31-49.

SCHERRE, M. M. P. *Reanálise da concordância nominal em português*. Rio de Janeiro: UFRJ, 1988. Tese (Doutorado em Linguística). Faculdade de Letras (UFRJ), Rio de Janeiro, 1988.

_____. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: OLIVEIRA, G. M. de; SCHERRE, M. (org.). *Padrões sociolingüísticos*. Rio de Janeiro: Tempo Brasileiro, 1996, p. 85-118.

_____. Concordância nominal e funcionalismo. *Revista Alfa*, v.41, p.181-206, 1997.

SCHERRE, M.; NARO, A. J. A concordância de número no português do Brasil: um caso de variação inerente. In: HORA, D. da (org.). *Diversidade linguística no Brasil*. João Pessoa: Ideia, 1997, p. 93-114.

VALLE, C. R. M.; GÖRSKI, E. M. Por um tratamento multidimensional da variação estilística na entrevista sociolingüística. In: GÖRSKI, E. M.; COELHO, I. L.; NUNES DE SOUZA, C. M. (org.). *Variação estilística: reflexões teórico-metodológicas e propostas de análise*. Coleção Linguística. v.3. Florianópolis: Insular, 2014, p. 93-121.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança linguística*. São Paulo: Parábola Editorial, 2006 [1968].



VALOR DIMINUTIVO DO SUFIXO -ÃO EM PORTUGUÊS

DIMINUTIVE VALUE OF THE SUFFIX -ÃO IN PORTUGUESE

Graça Rio-Torto¹

RESUMO

Neste estudo descreve-se o uso do valor diminutivo do sufixo *-ão* em português, seja na variante europeia, em cuja língua não padrão se encontra mais atestado, seja na variante brasileira, na qual tem uso diminuto. Uma vez que a identidade de uma língua não se confina à língua comum, mas inclui as suas manifestações diatópicas, diastráticas, diafásicas e diacrónicas, foram compulsadas fontes dialectais do português europeu que abonam a ocorrência do sufixo com valor diminutivo (cf. *pontão* ‘pequena ponte sobre um ribeiro, pequeno viaducto em estradas’; *ribeirão* ‘curso de água menor que um rio e maior que um riacho’), o qual está igualmente presente em outras línguas românicas, e analisam-se as condições estruturais e históricas da ocorrência deste padrão românico no idioma lusitano. Embora o valor semântico de diminuição de *-ão* esteja menos representado, sendo claramente periférico na atualidade, ele tem reflexos relevantes nas implicações que a duplicidade de ‘diminuição’ e ‘aumentação’ comporta para o estatuto dos avaliativos e para um dos fundamentos maiores da morfologia construcional e cognitivista: o pareamento entre estrutura semântica e estrutura fonológica. Propõe-se, então, uma solução compatível com este princípio e com a dualidade semântica de *-ão*.

PALAVRAS-CHAVE: Derivação; Sufixos avaliativos; Sufixos Diminutivos; Formação de palavras; Língua Portuguesa.

ABSTRACT

This study describes the use of the diminutive value of the suffix *-ão* in Portuguese, either in the European variant, in whose non-standard language is more attested, or in the Brazilian variant, in which it has small use. As the identity of a language is not confined to the common language, but includes its diatopic, diastractic, diaphasic and diachronic manifestations, dialectal sources of the European Portuguese were explored, since they support the occurrence of the suffix with a diminutive value (cf. *pontão* ‘little bridge [ponte] over a brook, small viaduct on roads’; *ribeirão* ‘water course smaller than a river and greater than a stream’), which is also present in other Romance languages. The structural and historical conditions of the occurrence of this Romance pattern in the Lusitanian language are analyzed. Although diminutive value of *-ão* is less represented, being clearly peripheral today, it has relevant repercussions on the implications that the duplicity of ‘diminution’ and ‘augmentation’ have for the status of evaluative suffixes and for one of major foundations of construcional and cognitive morphology: the pairing between semantic structure and phonological structure. A solution compatible with this principle and the semantic duality of *-ão* will be proposed.

KEYWORDS: Derivation; Evaluative suffixes; Diminutive suffixes; Word formation; Portuguese Language.

1 Professora Catedrática do Departamento de Línguas, Literaturas e Culturas/Membro do Celga-Iltec, da Faculdade de Letras (FLUC) da Universidade de Coimbra (UC). gracart@gmail.com.

Introdução

Nas gramáticas de referência do português, como as de Cunha e Cintra (1984: 91-92) e Bechara (2004: 361), o sufixo *-ão* é descrito como um aumentativo, não havendo qualquer referência ao seu valor diminutivo. Rocha Lima (1972) e Azeredo (2010) também não lhe fazem menção. O mesmo se aplica a Soares Barbosa (1822), Michaëlis de Vasconcelos (1914), Basílio (1999, 2004), Kehdi (1992) e Sandmann (1989).

O valor aumentativo de *-ão* é, em regra, o único assinalado, não sendo sequer mencionado o seu sentido diminutivo, por aquele ser de longe o semantismo mais representado e disponível.

Em todo o caso, a realidade da língua portuguesa, usada em Portugal ou no Brasil, revela-se mais complexa, pois em alguns casos, na língua comum e, sobretudo, na linguagem do mundo rural de falantes menos alfabetizados, em várias zonas dialectais de Portugal, persiste o uso de *-ão* com valor diminutivo, o que conduz a uma pluralidade semântica do sufixo e a uma não univocidade forma-sentido, com consequências teóricas não despiciendas para a morfologia derivacional e para o estatuto do sufixo.

Com efeito, a dificuldade de atribuição de um só significado a cada sufixo, documentada pela presença de mais do que um semantismo associado à mesma forma, derroga a relação de pareamento entre forma e significado (Aronoff; Fudeman 2005: 38-40), fundacional da morfologia construcional e da Linguística Cognitiva (Langacker 1987; Bybee 2010), em favor de uma relação plurívoca.

Pressupostos teóricos e metodologia

O quadro teórico que espalda esta reflexão é multidimensional, envolvendo o pensamento de nomes de referência dos estudos morfológicos e lexicais, como Aronoff; Fudeman (2005), Basílio (1999, 2004), Booij (2007, 2010), Gonçalves (2016), Nunes (1989), Michaëlis de Vasconcelos (1914), Rio-Torto (1993, 2016). A língua é encarada no seu uso, nas suas dimensões construcional, mental, cognitiva, sociocultural e pragmática, na sua variação diatópica, diafásica e diastrática, e nas suas sincronias e diacronias.

Dados de diferentes sincronias e de diferentes universos sociodialectais podem contribuir para clarificar o estatuto e o semantismo de certos afixos, como o dos avaliativos, mormente quando alguns destes são propensos a uma acentuada polissemia.

Foram escrutinadas fontes lexicográficas, representativas da realidade brasileira e portuguesa, maioritariamente do último quartel do século XX, de uso comum e com abrangências e extensão diversas, tais como o *Dicionário da língua portuguesa* da Porto Editora, o *Dicionário*

da *Língua Portuguesa* da Domingos Barreira, o *Grande dicionário da Língua Portuguesa*, de Cândido de Figueiredo, o *Dicionário Priberam da Língua Portuguesa* [em linha], 2008-2021, o *Novo dicionário da língua portuguesa*, de Aurélio Ferreira (2ª edição), e o *Dicionário eletrônico Houaiss da Língua Portuguesa*.

As fontes dialectais por nós compulsadas, muitas das quais inéditas, como (i) teses de Licenciatura [então de cinco anos] das Faculdades de Letras da Universidade de Coimbra e da Universidade de Lisboa e (ii) Relatórios do *Inquérito Linguístico Boléo*, consultáveis na Faculdade de Letras de Coimbra, testemunham o uso de -ão com valor diminutivo, ainda que não muito produtivo e/ou disponível, sobretudo em variedades mais ‘rurais’ do português europeu. Todavia, alguns dos nomes derivados em -ão com valor diminutivo são comuns a todo o universo pluricontinental de língua portuguesa, o que se deve à sua longevidade na história da língua (*cordão, estradão, pontão*, todos atestados desde o século XIII).

O sufixo -ão: origem e manifestações do valor diminutivo

Em várias línguas românicas, entre as quais a portuguesa, a espanhola, a italiana, a romena, o valor mais atestado dos representantes de -ONE é o aumentativo ou intensivo. Mas o francês, o provençal, o siciliano, o calabrês, o catalão, o aragonês e o galego dispõem ainda do valor diminutivo (DIEZ, 1874, p. 315-317; HASSELROT, 1957, p. 225), que também existe, ainda que escassamente representado, no português. Diez considera que é natural que tenha sido o francês a língua que mais diretamente contribuiu para o aparecimento de -ão diminutivo em português, como o atesta *cordão*, certamente do francês *cordon*.

Segundo Hasselrot (1957), -ONE era um sufixo raro na literatura antiga. Deve ter sido contemporâneo de -ITTU, mas teria sido usado apenas na linguagem oral. Qualquer que seja a sua gênese, a preservação do valor diminutivo de -ão em Portugal, como em terras de língua aragonesa, catalã, franco-provençal, siciliana e calabresa, aponta para um traço românico presente em várias geografias, por certo tributário de vários fenômenos de contacto linguístico em regiões tão distantes e ‘marginais’ quanto o litoral atlântico da Península Ibérica e o do sul de Itália (MAGNO, 1961), e de que há testemunhos inequívocos, nomeadamente no tocante às denominações de alfaias e utensílios agrícolas, como as denominações de medidas de capacidade e as de corte (PINTO, 1981, 1983). Os grandes manuais recentes de linguística românica (HARRIS; VINCENT, 1990; LEDGEWAY; MAIDEN, 2016, 2022) não fazem menção à disseminação e às razões da dispersão dos descendentes de -ONE com valor diminutivo, os quais sinalizam uma rede de intercomunicação lexical certamente muito antiga que, fruto de vicissitudes várias, se manteve como dominante em certas latitudes, e minoritária noutras.

O valor diminutivo de *-ão* é um valor não disponível no português coloquial contemporâneo, mas atestado em palavras recolhidas na linguagem popular de algumas variedades linguísticas mais conservadoras. Provavelmente, dos nomes de valor diminutivo abaixo arrolados, apenas *pontão* ‘pequena ponte sobre um ribeiro; pequeno viaduto em estradas’ e *estradão* ‘estrada ou caminho rústico e irregular, geralmente sem bermas definidas (*estradão em terra batida*)’ serão reconhecidos e usados com valor diminutivo na língua comum. No *Dicionário Priberam da Língua Portuguesa* [em linha], 2008-2021, <https://dicionario.priberam.org/> [consultado em 09-11-2021], de onde foram extraídas estas descrições, o primeiro semantismo abonado de *estradão* é o de ‘estrada grande ou larga’, confirmando assim a duplicidade de valores avaliativos associada a *-ão*.

Primitivamente este sufixo tinha por função explicitar, de forma neutra, a individualização ou a singularização de algo em relação ao que a base denota, mas progressivamente foi adquirindo os valores aumentativo e diminutivo. Como afirma o filólogo,

«ONE servait primitivement à individualiser; il s'attache à des thèmes verbaux ou nominaux pour désigner la personne qui accomplit une action avec une prédilection particulière, qui possède une qualité à un degré supérieur, qui se fait remarquer, attire parmi les autres l'attention par son occupation, par une particularité interne ou externe. Or cette signification fondamentale s'est diversement développée dans les différentes langues: -one est devenu dans presque tout le domaine franchement augmentatif, mais il a aussi élargi davantage encore sa propriété d'individualiser» (MEYER-LÜBKE, 1895, §456)².

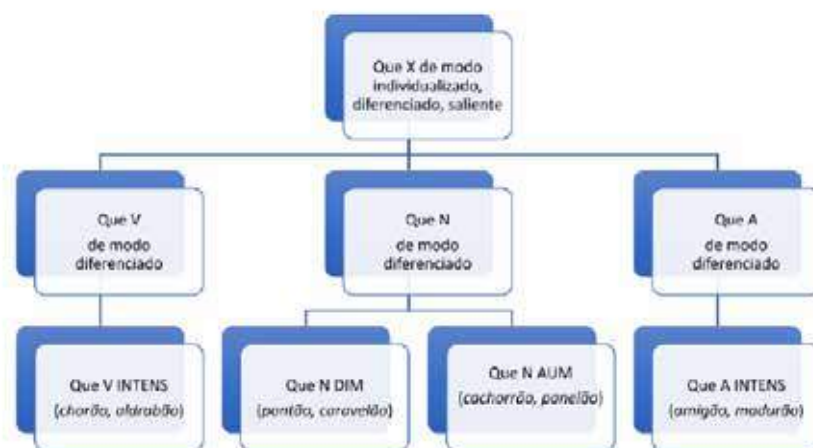
Assim, o valor matricial terá sido o de INDIVIDUALIZAÇÃO, através de uma dada propriedade saliente, que se cristalizou essencialmente no sentido aumentativo-intensivo, mas também no diminutivo. Na figura (1), o valor mais genérico, matricial e não subcategorizado de INDIVIDUALIZAÇÃO, de DIFERENCIAÇÃO é representado por “Que X de modo individualizado, diferenciado, saliente”; os valores de verbal, denominal e deadjetival são representados por “Que V de modo diferenciado”, “Que N de modo diferenciado”, “Que A de modo diferenciado” e manifestam-se, no uso concreto, por valores intensivos/aumentativos³ e,

2 Tradução nossa: «ONE servia primitivamente para individualizar; ele acopla-se a temas verbais e nominais para designar a pessoa que pratica uma ação com uma predileção particular, que possui uma qualidade em grau superior, que se faz notar, que atrai a atenção face aos outros pela sua ocupação, por uma particularidade interna ou externa. Ora, esta significação fundamental desenvolveu-se de modo diverso nas diferentes línguas românicas: -one tornou-se aumentativo em quase todo o domínio [românico], mas ele também alargou grandemente a sua propriedade de individualizar.»

3 Como Rio-Torto; Rodrigues (2016: 190-191 e 222-223) assinalam, este sufixo tem tendência a acoplar-se a bases não eruditas, formando nomes de evento culminado súbito e intenso, *como apertão, empurrão, encontrão, escaldão, rasgão*; no caso dos nomes de indivíduo (*berrão, chorão, reflão, saltão*), o indivíduo é individualizado pela intensidade e frequência com que efetua o evento. Ademais, há distribuição complementar entre os verbos que estão na origem de nomes de evento e dos de indivíduo, pois os verbos inacusativos e os de evento efetuado com força impulsiva apenas servem de base a nomes de evento, não a nomes de indivíduo.

apenas no caso dos nomes denominais, também pelo valor diminutivo. A Figura (1) sumariza as diferentes possibilidades combinatórias e respectivos valores semânticos de -ão.

Figura 1 – Do valor matricial aos valores de uso do sufixo -ão: esquemas construcionais e exemplos empíricos.



Fonte: elaboração própria

Para alguns, a emergência do valor diminutivo de -ão é parcialmente determinada pelo conteúdo da base a que o sufixo se anexa. Mas a coexistência de produtos nominais passíveis de uma leitura aumentativa e diminutiva atesta que o semantismo da base não tem influência sobre o valor do sufixo. Palavras como *albardão*, *alqueirão*, *caldeirão*, *casão*, *estradao*, *feirão*, *leirão*, *milhão*, *montão*, *peixão*, *pulgão*, *telhão* serão interpretadas, *a priori*, por um falante comum, como aumentativas; a sua leitura diminutiva ficará por certo circunscrita aos falantes de língua materna que já tenham tido acesso ao contacto com tais derivados, com esse valor mais antigo. Um caso paradigmático é o de *montão* que, nas edições dos anos 80 do *Dicionário da língua portuguesa* da Porto Editora ainda é descrito como ‘montículo, monte pequeno’, mas também como ‘conjunto de coisas empilhadas sem preocupação de ordem; acervo; pilha; aos montões ‘em grande quantidade’. Nas edições mais recentes do mesmo dicionário o valor diminutivo já não é abonado.

O sufixo -ão com valor diminutivo agrega-se fundamentalmente a bases marcadas pelo traço [-ANIMADO] (cf. *agulhão*; *albardão*; *batelão*; *caldeirão*; *camalhão*; *cangão*; *carreirão*; *casão*; *chavelhão*; *chicharões*; *cubatão*; *escadão*; *feirão*; *foução/foição*; *leirão*; *malotão*; *masseirão*; *milhão*; *montão*; *mosquetão*; *pedrões*; *pontelhão*; *regueirão*; *ribeirão*; *telhão*; *terrão*; *urtigão*) e pelo traço [+ANIMAL] (cf. *aguidão*; *aguidão*; *peixão*; *pintão*; *pulgão*). Não há registo de derivados construídos com base em radicais de nomes de ser humano, de nomes de qualidade ou de propriedade. Em português, e tanto quanto nos é dado saber, nas demais línguas românicas, o valor diminutivo do sufixo não está disponível para a derivação de nomes deverbais. Este sufixo também não é compatível com bases adjetivas. O seu espectro combinatório é,

pois, manifestamente diminuto face ao de *-ão* com valor aumentativo/intensivo que, sendo compaginável com bases nominais, adjetivais e verbais, tem muito mais condições estruturais para continuar disponível e produtivo, o que não acontece com *-ão* de valor diminutivo.

Alguns casos, como o de *mantão* ‘(ant.) espécie de capote curto’ (**Figueiredo 1981**), ‘espécie de capotilha’ (**Silva 1955**), ‘espécie de capa com capuz’ (**Machado 1977**) apresentam fronteiras ténues entre o valor diminutivo e o especializado. São aqui considerados os nomes que explicitamente possuam traços semânticos de diminuição, nas suas diversas modalidades e variantes ("tipo mais pequeno de X", em que X representa o denotado pela base nominal), ou seja, todos aqueles cuja significação pressupõe uma avaliação diminutiva que afeta as dimensões, a estatura, a idade, entre outras.

Na listagem de dados empíricos que se seguem (1-38), as referências lexicográficas correspondem aos seguintes dicionários do Brasil e de Portugal:

- **Costa; Melo 1984:** Costa, J. Á; Melo, A. S. Dicionário da língua portuguesa. 6.a ed. Porto, Porto Editora; 1984.
- **Ferreira 1987:** Ferreira, A. B. H. *Novo dicionário da língua portuguesa*. 2a edição. Rio de Janeiro, Nova Fronteira; 1987.
- **Figueiredo 1981:** Figueiredo, C. *Grande dicionário da língua portuguesa*. 2 vols. Lisboa, Bertrand. 16a edição, 1981.
- **Houaiss 2002:** Houaiss, A. Dicionário eletrônico Houaiss da língua portuguesa. Rio de Janeiro, Editora Objetiva; 2002.
- **Silva 1955:** Silva, F. J. Dicionário da língua portuguesa. 3.a ed. Porto, Domingos Barreira; 1955.

São os seguintes os derivados recolhidos nas fontes lexicográficas e dialectais compulsadas:

- 1) **adelhão** — calha de madeira/pequena caldeira, suspensa da adelha, com inclinação regulada por um cordel [...] (**Figueiredo 1981**); espécie de caldeira que conduz o cereal da adelha para o centro da mó (**Costa; Melo 1984**).
- 2) **agudião** — formiga alada, mais pequena que a agúdia (**Figueiredo 1981**).
- 3) **aguidão** — agúida pequena (**Figueiredo 1981**).
- 4) **agulhão** — pequena bússula de bordo (**Costa; Melo 1984; Figueiredo 1981**); (náut.) agulha-padrão; pedra pontiaguda submersa no leito dum rio (**Ferreira 1987**).
- 5) **albardão** — albarda pequena, no falar popular de Odeleite, Algarve (Cruz, 1991, p. 146).

- 6) **barranhão** — (ou *barrenhão*) recipiente de madeira onde se prepara a comida para os porcos; pequeno alguidar (**Costa; Melo 1984**); (prov.) pequeno barril portátil para vinho (**Silva 1955**); alguidar pequeno, não vidrado, na linguagem popular de Aldeia Nova de São Bento, Serpa, Beja (Seita 1944, p. 46); pequeno alguidar ou vasilha de barro grosseiro, semelhante ao alguidar, mas mais pequeno, onde comem os ganhões (Pombinho Jr. 1924, p. 74).
- 7) **batelão** — (bras. MT) canoa pequena (**Ferreira 1987**).
- 8) **cangão** — canga pequena, sem varandas ou molduras; registado em Venade, Caminha, Viana do Castelo (Loução 1931, p. 251).
- 9) **caravelão** — (ant.) antiga embarcação de vela, de porte inferior às vulgares nesse tempo e sem acabamento perfeito (**Costa; Melo 1984**); caravela pequena e rudimentar, muito utilizada no litoral brasileiro no início da colonização (**Ferreira 1987**); já atestada no *Tratado descritivo do Brasil (1587)* de Gabriel Soares de Sousa (cf. Primeira Parte: Roteiro geral da Costa brasílica, capítulo IX, p. 50; *ibidem*, p. 124).
- 10) **carreirão** — (prov. transm.) caminho estreito para carros; pequeno carreiro (**Figueiredo 1981**); pequeno caminho de carros; um carreiro (SANTOS, 1967, p. 243).
- 11) **casão** — pequena casa onde se arruma o material da pesca, também designada por *casino* na Ericeira, Mafra, Lisboa (Alves 1965, p. 205).
- 12) **chavelhão** — peça de madeira, mais pequena que a *chavelha*, que se enfia na cabeçalha, à frente da canga, registado em Vila Nova de Ourém, Santarém (Silva 1972, p. 277).
- 13) **chicharões** — feijões ‘chícharos’ pequenos, na linguagem popular da Beira Baixa (Dias 1962, p. 33).
- 14) **cobrão** — cobrelo ‘cobra pequena’ (**Costa; Melo 1984**).
- 15) **cubatão** — (bras. SP) pequena elevação no sopé de cordilheiras (**Ferreira 1987**).
- 16) **dobrão** — antiga moeda portuguesa de ouro, que valia 24 réis (**Figueiredo 1981**); (bras. do N e NE) moeda antiga de cobre, do valor de 40 réis (**Ferreira 1987**). Base: radical de dobra, nome de antiga moeda de ouro, que Santa Rosa Viterbo descreve, no *Elucidário* (vol. II, 1966, p. 199-200), como tendo valido entre 60 a 270 réis.
- 17) **escadão** — pequena escada; escadote (R.I.L.B., Castelo, Arnóia, Celorico de Bastos, Braga. 1952, p. 7).

- 18) **escotilhão** — escotilha pequena (**Silva 1955**); pequena escotilha (**Ferreira 1987**).
- 19) **feirão** — (lus.) pequena feira (**Ferreira 1987**); (prov. minh.) feira pequena (**Costa; Melo 1984**); feirinha; este conteúdo, que permite opor a *feira* (mensal) ao *feirão* (mercado semanal) ainda se preserva em Montalegre (Girão 1957, p. 55), na ‘Terra Quente’ transmontana (concelhos de Alfândega da Fé, Carrazeda de Ansiães, Macedo de Cavaleiros, Mirandela, Vila Flor) e em todo o Alto Minho (Pereira 1917, p. 245).
- 20) **foiçã** — foice pequena (**Costa; Melo 1984**).
- 21) **fouçã** — instrumento para cortar cereais, de gume serrilhado, que designa a FALX MESSORIA, e que é mais pequeno que a fouce (foice); está representado nos distritos da Guarda, Castelo Branco e Santarém (Pinto 1981, p. 132).
- 22) **leirão** — leira pequena (**Figueiredo 1981**); espaço de terreno cultivado menor que uma leira (**Figueiredo 1981**); uma das partes em que se divide a leira, no concelho de Figueiró dos Vinhos (Silva 1960, p. 253).
- 23) **masseirão** — masseira pequena para usos diversos, nomeadamente para servir alimentos aguados a animais domésticos (**Costa; Melo 1984**).
- 24) **milhão** — (ant.) milho miúdo (**Figueiredo 1981**); milho que não logrou atingir o crescimento pleno, pelo que é aproveitado como alimento para o gado (zonas rurais do concelho de Águeda, Aveiro); grão de milho (R.I.L.B., S. Martinho, Várzea da Serra, Tarouca, Viseu. 1965, p. 158).
- 25) **montão** — montículo, monte pequeno (**Costa; Melo 1984**); “*cavar de montão*” ‘fazer a cava, deixando a terra em pequenos montes’, no falar dos lavradores e pescadores do concelho de Olhão, Algarve (Macara 1964, p. 200).
- 26) **mosquetão** — (bras.) fuzil pequeno usado pelos soldados de cavalaria e de artilharia (**Ferreira 1987**); arma de cano curto, usada pela cavalaria e pela artilharia no Brasil (**Figueiredo 1981**). Mosquete ‘espingarda grossa e pesada’.
- 27) **narcejão** — ave migradora parecida com a galinhola, porém mais pequena (**Figueiredo 1981**).
- 28) **peixão** — goraz, quando juvenil (**Costa; Melo 1984**); (t. de Aveiro e Cascais) besugo pequeno (**Figueiredo 1981**); pequeno peixe de cor prateada, semelhante ao besugo, no concelho de Vila do Conde (Netto 1949, p. 130).
- 29) **pintão** — filho da galinhola; frango; pintalegrete (**Silva 1955**).
- 30) **pontelhão** — pequena ponte (**Silva 1955**).
- 31) **pontilhão** — pequena ponte de vão total inferior a 10 metros (**Ferreira 1987**).

- 32) **quartão** — vasilha de barro para água, menor que a *quarta* mas maior que a bilha, em Turquel, Leiria (Pinto, 1981, p. 176-177); nome de medida de líquidos que, de um modo geral, equivale a um quarto de almude, e que está representado nos distritos de Aveiro, Coimbra, Leiria, Santarém, Castelo Branco, Portalegre (Pinto 1983, p. 166 e p. 176-177).
- 33) **quarteirão** — medida de líquidos, correspondente à quarta parte do *quartilho*; o seu valor oscila entre um quarto e um oitavo de litro; medida de capacidade equivalente à quarta parte do *quarteiro* (Pinto 1983, p. 25 n. 4 e p. 178-179).
- 34) **regueirão** — pequena corrente de água, registado em Bruçó, concelho de Mogadouro e Lagoaça, e no concelho de Freixo de Espada à Cinta, Bragança (Santos 1967, p. 243).
- 35) **ribeirão** — curso de água menor que um rio e maior que um riacho (**Ferreira 1987**); já registado como "nomen unitatis" em Trás-os-Montes (Santos 1967, p. 243).
- 36) **rodilhão** — pequena roda, em zorras e carros de mão (**Figueiredo 1981**).
- 37) **telhão** — telha prensada (**Silva 1955**) ou pequena (registado em Águeda, Aveiro); (reg. da Bairrada, Beira Litoral) pedaço de telha partida, ou um caco de telha (**Figueiredo 1981**).
- 38) **urtigão** — urtiga miúda, que se aproveita para os perús enquanto pequenos, no Parâmio, Bragança (Fernandes 1961, p. 111).

Embora muitos dos derivados diminutivos em *-ão* sejam provenientes de universos lexicais regionais e/ou da linguagem popular, o sufixo *-ão* com valor diminutivo não é um operador exclusivo da linguagem popular. Ainda que se trate de um instrumento derivacional não comum e de uso predominantemente popular e não urbano, a sua ocorrência não se restringe a variedades idiomáticas específicas, pelo que ele deve ser encarado como um sufixo pouco produtivo, não disponível no português contemporâneo, que está tanto mais documentado quanto mais se recua na linguagem ‘popular’ do século XX.

Das quase quatro dezenas de nomes arrolados, estão abonados nos dicionários brasileiros compulsados oito derivados: *agulhão*, *batelão* (bras. MT), *caravelão*, *cubatão* (bras. SP), *dobrão* (bras. do N e NE), *feirão*, *mosquetão*, *pontilhão*. No *Dicionário eletrônico Houaiss da Língua Portuguesa*, o derivado *batelão* tem valor diminutivo e também aumentativo e *feirão* é descrito como Lusitanismo. *Agulhão* 'Pedra aguda, submersa no leito de um rio', denota ainda '[Náutica] Grande agulha de marear', sentido também registado nos dicionários lusitanos.

Em Portugal, o recurso a este sufixo está igualmente atestado na formação de hidrónimos e de "topónimos correlatos", isto é, de nomes de pequenas localidades, casais, aldeolas ou

lugarejos cuja denominação representa um diminutivo formado a partir de um outro topónimo, mais antigo e de uma localidade mais significativa.

Como hidrónimo pode mencionar-se *Mondegão*, nome de um pequeno curso de água que desce da Serra da Estrela e desagua em ribeiros que vão dar ao Mondego (GIRÃO, 1955, p. 90), e que significa, em última análise, um pequeno Mondego.

São produtos deste tipo: *Cabeção*, topónimo muito frequente no centro e no sul do país; *Castelão e Castelões*, topónimos diminutivos largamente representados em diversas regiões de Portugal; *Caldeirão*, nome da lagoa vulcânica da mais pequena ilha açoreana, o Corvo, contrastando com as caldeiras que, nas outras ilhas, designam crateras maiores (GIRÃO, 1957, p. 55-57); *Albardão*, nome de freguesia de São Manços, concelho de Évora; *Paranhão*; *Picão*; *Valegões* (PIEL, 1940, 1947); *Monchicão*, pequeno lugar da freguesia de Alferce, concelho de Monchique.

O quadro (1) sumaria os dados coligidos.

Quadro 1 – Topónimos e hidrónimos portugueses em *-ão* com valor diminutivo.

Topónimos correlatos		Hidrónimo
Albardão	Monchicão	
Cabeção	Paranhão	Mondegão
Caldeirão	Picão	
Castelão	Valegões	
Castelões		

Fonte: elaboração própria.

No Brasil, se tomarmos como amostra a que foi recolhida por Tavares (2019, p. 132) na região centro-norte de Mato Grosso do Sul, os topónimos derivados em *-ão* que foram identificados têm todos semantismo aumentativo: *Areão, Baixadão, Brejão, Buracão, Canastrão, Corixão, Grotão, Lagoão, Lajeado, Manecão, Perdigão, Pindaibão, Poção, Taperão*. Muitas outras possibilidades de formação estavam em jogo (fitónimos, hidrónimos, zoónimos) que poderiam recorrer ao sufixo *-ão* com valor diminutivo, caso este fosse usado. Mas a sua ausência atesta a sua indisponibilidade no contexto em pauta.

Que consequências para o estatuto de *-ão* e para a teoria morfológica?

A descrição do modo de funcionamento do sufixo *-ão* na língua portuguesa do último século, usada no Brasil e em Portugal, beneficia com o fato de se ter em conta a operacionalização do sufixo por falantes de universos socioculturais e dialectais de esferas menos ‘padrão’, nomeadamente no âmbito do léxico ‘popular’ do mundo rural lusitano e no tocante a alguns dos seus lexemas mais antigos, os quais ajudam a compreender a globalidade dos usos do sufixo na sincronia do século XX. A realidade geo-sociolectal mais especiosa, tipicamente mais

diferenciada que a da língua comum, proporciona uma abrangência de situações, contextos e valores que alargam o mapa de relações, funções e valores de uso de um dado sufixo, e assim acontece também no caso do operador aqui em apreço: *-ão*.

Como vários estudiosos assinalaram (ROSA, 1982; GONÇALVES et al., 2009), o sufixo *-ão* é, antes de tudo, um sufixo aumentativo e intensivo, seja em nomes e em adjetivos isocategoriais (RIO-TORTO, 1993), ou seja, nomes derivados de nomes, como *grupão, janelão, jantarão, trabalhão*, e adjetivos derivados de adjetivos, como *amigão, gostosão, madurão*, seja também em nomes deverbais, os quais denotam ‘agente/aquele/aquilo que V’ (*aldrabão, mandão, fujão*), ‘instrumento com que se V’ (*esfregão, picão*), ‘evento, ação/processo de V’ (*arrastão, esticão, puxão*). Além deste semantismo, os deverbais apresentam também um valor de intensidade, mais saliente nos agentivos e nos eventivos que nos instrumentais, e que foi interpretado (RIO-TORTO, 2020) como atribuído por herança do valor mais prototípico do sufixo, o valor aumentativo-intensivo. Sem descartar esta possibilidade, a gênese e a história do sufixo apontam para que, em complemento, a sua identidade tenha passado originalmente pela INDIVIDUALIZAÇÃO/DIFERENCIAÇÃO da entidade/do evento em jogo, através de uma propriedade saliente, que se terá fixado num sentido mais aumentativo-intensivo em algumas línguas, e num sentido diminutivo noutras. Na língua portuguesa hodierna, ambos os valores podem conviver, mas não se sobrepõem nas mesmas circunstâncias comunicativas.

Como já afirmado, o sentido diminutivo é claramente periférico face ao mais central, que é o aumentativo-intensivo. Não sendo desejável estabelecer uma homonímia entre dois sufixos *-ão*, um diminutivo e um aumentativo-intensivo, importa saber como incorporar a polissemia do sufixo na relação de pareamento entre forma e significado, há muito postulada como crucial pela teoria linguística. Um parêntesis para explicitar que não consideramos estar perante um caso de heterossemia, tal como descrita por Lichtenberk (1991), pelo fato de os dois valores semânticos, que estão relacionados historicamente — sendo, portanto, espaldados por uma comum origem e matriz (etimológica e semântica) —, não pertencerem a categorias morfossintáticas diversas, não desempenharem funções pragmáticas diversas, e não serem resultantes de um processo de gramaticalização.

Assim, a partir do esquema construcional mais genérico e abrangente ‘que X de modo individualizado, diferenciado, saliente’, em que X representa aquilo que a base nominal, adjetival ou verbal denota, são gerados esquemas mais específicos ‘que N de modo diferenciado’, ‘que V de modo diferenciado’, ‘que A de modo diferenciado’, em que as manifestações dos modos diferenciados se articulam com a natureza nominal, verbal ou adjetival da base, manifestando-se por *aumentação, intensidade, diminuição*. O valor diminutivo apenas está presente em nomes denominais.

A coabitação de semantismos aumentativo-intensivos e diminutivos, como a de aumentativo e depreciativo ou diminutivo e depreciativo, é frequente em vários dos operadores avaliativos do português (*-inh-*, *-ec-*, *-alh-*, entre outros), pelo que a sua coexistência nos derivados em *-ão* não representa uma anomalia. Este sufixo revela-se, como outros avaliativos, capaz de albergar uma polissemia de manifestações de avaliação que está inscrita na própria natureza da avaliação, seja qualitativa, quantitativa e/ou intersubjetiva.

No caso de *-ão*, as significações que o sufixo ganhou, face às matriciais, não alteraram o significado do molde original, antes o modalizaram sob a forma de extensões em nada arbitrárias relativamente ao valor primordial. Também no caso de *-inh-*, o valor hedónico (*arrozinho*, *banhinho*, *carninha*, *sol(z)inho*), afetivo (*filhinho*, *mãezinha*, *sapatinho*), o de proximidade e empatia intersubjetiva (*adeusinho*, *um cafezinho*, *um favorzinho*, *um minutinho*) inscrevem-se igualmente nas possibilidades funcionais e pragmáticas de um diminutivo (na linha de que ‘o que é pequeno tem pouco valor, é depreciável, o que é pequeno é apreciável, estimável, empático’). A coexistência de sinais antagónicos, num sentido majorativo ou minorativo, apreciativo ou depreciativo, inscreve-se, portanto, na radialidade centro/protótipo-periferia(s) que sustenta o ‘caminho’ da individualização/diferenciação até à avaliação (aumentativa/diminutiva) e que faz parte do DNA dos avaliativos. Essa mesma radialidade centro/protótipo-periferia(s) está também patente de forma emblemática nos sufixos avaliativos que acolhem, com ponderações distintas, os valores aumentativo e diminutivo, como acontece com *-ão*, *-inh-* e *-alh-*.

Uma vez mais se atesta que os significados são estruturados, relativamente abertos e subjetivos. Um operador pode ter um significado primário, mas o seu significado global pode incluir não só esse semantismo primário e central, como outros sentidos mais e menos centrais ou mais e menos periféricos, em função dos contextos situacionais em que é usado. No caso de *-ão*, o pareamento entre forma e sentido não é posto em causa, se assumirmos que o valor do operador é o avaliativo, e que este se pode manifestar num sentido diminutivo ou aumentativo.

Ao compaginar valor aumentativo-intensivo e diminutivo no sufixo *-ão*, a língua portuguesa fica assim, duplamente irmanada com as demais línguas românicas, como o francês, ou o catalão, línguas que dispõem ainda do valor diminutivo, e também comunga com o castelhano, o italiano, o romeno, o valor aumentativo-intensivo. Mas no conspecto das línguas românicas, a língua portuguesa distingue-se das demais por acumular, ainda que com pesos diferentes, os valores aumentativo e diminutivo do descendente de *-ONE*, como o quadro seguinte visualiza.

Quadro 2 – Valor aumentativo e diminutivo dos descendentes românicos de -ONE

	Valor AUMENTATIVO	Valor DIMINUTIVO
Português -ÃO	Amigão, calorão, churrascão, cobardão, tempão, toalhão	Cordão, estradão, pontão
Espanhol -ÓN	Cucharón, cobardón, hombrón, fortunón	
Italiano -ONE	Casone, gattone, palazzone ragazzone, successone	
Francês -ON	-	Aiglon, botillon, cordon, raton
Catalão -Ó	-	Animaló, calentó, carretó

Fonte: elaboração própria

Uma vez mais, o conhecimento da história e da expressão multilinguística dos descendentes de um mesmo étimo na sua família de línguas, facilita a compreensão dos atuais valores de um operador (no caso sufixal), sejam os mais centrais, sejam os mais periféricos. Uns e outros são relevantes para o ‘mapeamento’ do funcionamento do sufixo, na sua diversidade semântica e na sua diferenciação transcontinental.

Considerações finais

No português europeu e, em menor escala, no português do Brasil, o sufixo -ão funciona como operador diminutivo, estando presente em alguns nomes do léxico comum, como *estradao*, *pontão*, e em nomes do léxico mais específico do mundo rural, brasileiro (v.g. *cubatão*, *mosquetão*) ou lusitano (*agulhão*; *albardão*; *batelão*; *caldeirão*; *camalhão*; *cangão*; *carreirão*; *casão*; *chavelhão*; *chicharões*; *escadão*; *feirão*; *foução/foição*; *leirão*; *malotão*; *masseirão*; *pedrões*; *peixão*; *pintão*; *pulgão*; *pontelhão*; *regueirão*; *ribeirão*; *telhão*; *terrão*; *urtigão*). O valor diminutivo é raramente mencionado, talvez por ser menos típico e bem menos usado que o aumentativo e intensivo, de longe o mais representativo do sufixo. Por isso, para um falante comum, o mais provável é que os nomes sufixados em -ão acima arrolados sejam interpretados como aumentativos, correspondendo a ‘x [o denotado pela base] de grandes dimensões’. Tal valor diminutivo, que é comum a outras línguas românicas, encontra-se atestado em lexemas denotadores de realidades do mundo rural de Portugal e do Brasil, estando também presente em palavras já antigas na língua, como *caravelão*, *dobrão*, *mosquetão*, *ou quartão* e *quarteirão* ‘denominações de medidas de líquidos’.

A coexistência do valor diminutivo e do valor aumentativo, ainda que com pesos diferentes e com usos contextualmente distintos, é uma propriedade que afeta outros sufixos avaliativos, e que se enraíza no semantismo do próprio sufixo -ONE. O valor matricial deste terá sido o de INDIVIDUALIZAÇÃO/DIFERENCIAÇÃO, através de uma dada propriedade saliente que, na língua portuguesa, se cristalizou essencialmente no sentido aumentativo-intensivo, mas também no diminutivo. Este valor diminutivo manifesta-se nos nomes denominais; o valor aumentativo-

intensivo está presente em nomes denominais, deverbais e em adjetivos. Assim, *-ão* comporta-se como outros avaliativos (v.g. *-inh-*, *-ec-*) que, ao lado de um valor mais representado, como o diminutivo de *-inh-* (*dedinho*, *livrinho*) ou o depreciativo de *-ec-* (*lojeca*), acomodam um valor intensivo (*certinho*, *pertinho*) ou um valor apreciativo (*soneca*, *monstreco*). No caso de *-ão* os valores em coexistência são o diminutivo e o aumentativo, com claro ganho de uso deste. O pareamento entre forma e sentido não fica posto em causa, se assumirmos que o valor do operador é o avaliativo, e que este se pode manifestar num sentido diminutivo ou aumentativo.

Para compreender o valor diminutivo de *-ão* torna-se, pois, necessário convocar conhecimentos históricos, que amiúde nos trazem luz sobre os usos mais próximos do operador no transcurso da língua. As abordagens a-históricas da morfologia e da semântica dos operadores afixais podem ocultar factos e realidades que ajudam a explicar satisfatoriamente os usos atuais dos mesmos. A utilidade do concurso da morfologia histórica e diacrónica torna-se, neste caso de *-ão*, inquestionável.

Ademais, ao permitir a coexistência dos valores aumentativo-intensivo e diminutivo no sufixo *-ão*, a língua portuguesa comunga das duas grandes redes idiomáticas que se formam no interior das línguas românicas: (i) a que inclui o francês, o provençal, o siciliano, o calabrês, o catalão, o aragonês e o galego, línguas cujos representantes de *-ONE* dispõem ainda do valor diminutivo; e (ii) a que congrega o castelhano, o italiano, o romeno, em que os descendentes de *-ONE* têm valor aumentativo-intensivo. Falta conhecer a fundo as dinâmicas dos fluxos lexicais que sustentam historicamente estas duas facetas do universo românico. Ao mesmo tempo, a língua portuguesa singulariza-se pela coexistência dos valores aumentativo-intensivo e diminutivo do sufixo *-ão*, o que é uma peculiaridade já assinalada (RIO-TORTO, 2022). Uma vez mais, o conhecimento da história e da expressão multilinguística dos descendentes de um mesmo étimo na sua família de línguas propicia uma compreensão mais holística dos atuais valores de um operador sufixal, sejam os mais centrais, sejam os mais periféricos. Uns e outros são relevantes para o mapeamento do funcionamento do sufixo, na sua diferenciação intrassistémica e diassistémica, em razão da rede de contactos desde sempre mantidos entre as línguas românicas de diferentes latitudes.

Referências

ALVES, J. L. M. R. L. *A linguagem dos pescadores da Ericeira*. Lisboa, Junta Distrital de Lisboa, 1965.

ARONOFF, M.; FUEDEMAN, K. *What Is Morphology?* Oxford: Blackwell, 2005.

AZEREDO, J. C.de. *Gramática Houaiss da língua portuguesa*. São Paulo: Publifolha, 2010.

BASÍLIO, M. *Formação e classes de palavras no PB*. São Paulo: Contexto, 2004.

BASÍLIO, M. *Teoria Lexical*. 6. ed. São Paulo: Ática, 1999.

Diadorim, Rio de Janeiro, vol. 24, número 1, p. 233 - 250, 2022.

BECHARA, E. *Moderna gramática portuguesa*. 37 ed. rev. e ampliada. Rio de Janeiro: Lucerna, 2004.

BOOIJ, G. *The Grammar of Words: An Introduction to Linguistic Morphology*. 2nd edition. Oxford: Oxford University Press, 2007.

BOOIJ, G. *Construction Morphology*. Oxford: Oxford University Press, 2010.

BYBEE, J. *Language, usage and cognition*. Cambridge: Cambridge University Press, 2010.

COSTA, J. A.; MELO, A. S. e. *Dicionário da língua portuguesa*. 6ª edição, corrigida e aumentada. Porto: Porto Editora, 1984.

CUNHA, A. G. da. *Dicionário etimológico Nova Fronteira da língua portuguesa*. 2ª edição, revista e acrescida de um suplemento. Rio de Janeiro: Editora Nova Fronteira, 1986.

CUNHA, C.; CINTRA, L. F. L. *Nova gramática do português contemporâneo*. 2 ed. Lisboa: Edições João Sá da Costa, 1984.

CRUZ, M. L. S. da. *O falar de Odeleite*. Lisboa, Fundação Calouste Gulbenkian; Instituto Nacional de Investigação Científica, 1991.

DIAS, J. L. *A linguagem popular da Beira Baixa*. Lisboa: Editorial Império, 1962.

Dicionário Priberam da Língua Portuguesa [em linha], 2008-2021. <https://dicionario.priberam.org/> [consultado em 09-11-2021].

DIEZ, F. *Grammaire des langues romanes*. Tome II. Paris, A. Franck, 1874.

FERNANDES, H. da P. *O Parâmio. Contribuição para o estudo da linguagem e etnografia da região bragançana*. Dissertação de Licenciatura (inérita). Faculdade de Letras. Universidade de Coimbra. Coimbra. 1961.

FERREIRA, A. B. de H. *Novo dicionário da língua portuguesa*. Rio de Janeiro: Nova Fronteira, 1975; 2ª edição. 1987.

FIGUEIREDO, C. de. *Grande dicionário da língua portuguesa*. 2 vols. Lisboa: Bertrand. 1º vol., 16ª edição, 1981; 2º vol., 15ª edição, 1982.

GIRÃO, A. de A. Mondego, *Mondeguinho e Mondegão*. In: Boletim do Centro de Estudos Geográficos, nº 10-11, p. 90-91, 1955.

GIRÃO, A. de A. *O sufixo -ão como diminutivo toponímico*. In: Boletim do Centro de Estudos Geográficos, nº 14-15, p. 55-57, 1957.

GONÇALVES, C. A. V. et al. *Para uma estrutura radial das construções X-ão do português do Brasil*. In: ALMEIDA, M. L. L. et al. (org.). *Linguística Cognitiva em foco: morfologia e semântica*. Rio de Janeiro: Publit, 2009. p. 141-156.

GONÇALVES, C. A. V.; ALMEIDA, M. L. L. de. Morfologia construcional: principais ideias, aplicação ao português e extensões necessárias. *Alfa*, São Paulo, v. 58 nº 1, p. 165-193, 2014.

GONÇALVES, C. A. *Atuais tendências em formação de palavras*. São Paulo: Contexto, 2016.

HARRIS; M; VINCENT, N. (ed.). *The Romance Languages*. Oxford: Oxford University Press. 1990.

HASSELROT, B. *Études sur la formation diminutive dans les langues romanes*. Upsala: Acta Universitatis Upsaliensis, 1957.

HOUAISS, A. *Dicionário eletrônico Houaiss da Língua Portuguesa*. Rio de Janeiro: Editora Objetiva, 2002.

KEHDI, V. *Formação de palavras em português*. São Paulo: Ática, 1992.

LANGACKER, R. W. *Foundations of cognitive grammar*. Vol. I: theoretical prerequisites. Stanford: University Press, 1987.

LEDGEWAY, A.; MAIDEN, M. (ed). *The Oxford Guide to the Romance Languages*. Oxford: Oxford University Press. 2016.

LEDGEWAY, A.; MAIDEN, M. (ed). *The Cambridge Handbook of Romance Linguistics*. Cambridge: Cambridge University Press. 2022.

LICHTENBERK, F. On the gradualness of grammaticalization. In: TRAUGOTT, E.; HEINE, B. (ed.). *Approaches to grammaticalization*. Amsterdam: John Benjamins Publishing Company, 1991. v.1, p. 37-80.

LOUÇÃO, J. L. L. Lexicografia das margens do Minho. In: *Revista Lusitana*, vol. XXIX, p. 246-276, 1931.

MACARA, M. C. C. *O falar dos lavradores e pescadores do concelho de Olhão. Contribuição para o estudo linguístico, etnográfico e folclórico*. Dissertação (Licenciatura em Filologia Românica). Faculdade de Letras. Universidade de Coimbra. Coimbra. 1964.

MACHADO, J. P. *Dicionário etimológico da língua portuguesa*. 3ª edição, 5 vols. Lisboa, Livros Horizonte, 1977.

MAGNO, L. M. dos S. *Áreas lexicais em Portugal e na Itália*. Dissertação de Licenciatura. Coimbra. Separata da Revista Portuguesa de Filologia, vol. XI. 1961.

MEYER-LÜBKE, W. *Grammaire des langues romanes* (traduction par Auguste Doutrepoint e Geroges Doutrepoint). Tome II: Morphologie. Paris: H. Welter Éditeur, 1895.

MICHAËLIS DE VASCONCELOS, C. *Lições de filologia portuguesa (segundo as preleções feitas aos cursos de 1911-1912 e 1912-1913)*. Lisboa: Dinalivro, 1914.

NETTO, M. T. de M. L. A linguagem dos pescadores e lavradores do concelho de Vila do Conde. Separata da *Revista Portuguesa de Filologia*, vol. I, p. 59-151 e vol. II, p. 122-187, Coimbra, 1949.

NUNES, J. J. *Compêndio de gramática histórica portuguesa - Fonética e Morfologia*. 9ª edição, Lisboa: Livraria Clássica Editora, 1989.

PEREIRA, F. A. Glossário dialectológico do concelho de Arcos de Valdevez (Alto-Minho). In: *Revista Lusitana*, vol. XX, p. 239-256, 1917.

PIEL, J. M. A formação dos nomes de lugares e de instrumentos em português. In: *Boletim de Filologia*, tomo VII, fasc. 1, p. 31-47, 1940.

PIEL, J. M. Nomes de lugar referentes ao relevo e ao aspecto geral do solo (capítulo de uma toponímia galego-portuguesa). Separata da *Revista Portuguesa de Filologia*, vol. I, tomo I, p. 153-197, 1947.

PINTO, A. A. *Áreas linguístico-etnográficas de alfaias agrícolas de corte*. Sep. de: Biblos (Coimbra. Faculdade de Letras), vol. 57, p. 97-163, 1981.

PINTO, A. A. *Isoléxicas portuguesas: antigas medidas de capacidade*. Coimbra: Universidade de Coimbra. Faculdade de Letras. Instituto de Língua e Literatura Portuguesas. Sep. de *Revista Portuguesa de Filologia*, vol. XVIII, 1983.

POMBINHO Jr., J. A. Vocabulário alentejano (subsídios para o léxico português). In: *Revista Lusitana*, vol. XXVI, p. 68-83, 1924.

Relatórios do Inquérito Linguístico Boléo. Faculdade de Letras de Coimbra. [R.I.L.B.]

RIO-TORTO, G. *Formação de palavras em português. Aspetos da construção de avaliativos*. Dissertação de Doutoramento em Linguística Portuguesa. Coimbra: Universidade de Coimbra, 977 p. + I-VI. Reprodução da Secção de Textos da Faculdade de Letras de Coimbra. 1993. <http://hdl.handle.net/10316/24452>.

RIO-TORTO, G. Formação de avaliativos. In: RIO-TORTO, G. et al., *Gramática derivacional do português*. Coimbra: Imprensa da Universidade de Coimbra, 2016, p. 357-389.

RIO-TORTO, G. Derivação. In: RAPOSO, E. P. et al., *Gramática do Português*. Lisboa: Fundação Calouste Gulbenkian, 2020, p. 215-336.

RIO-TORTO, G. Peculiarities of Portuguese Word-Formation. In: M. Loporcaro (Ed.), *Oxford Encyclopedia of Romance Linguistics* (part of *Oxford Research Encyclopedia of Linguistics* ed. by M. Aronoff). Oxford University Press, 2022. <https://doi.org/10.1093/acrefore/9780199384655.013.687>.

RIO-TORTO, G.; RODRIGUES, A. S. Formação de nomes. RIO-TORTO, G. et al., *Gramática derivacional do português*. Coimbra: Imprensa da Universidade de Coimbra, 2016, p. 135-240.

- ROCHALIMA, C. H. *Gramática Normativa da Língua Portuguesa*. Rio de Janeiro, J. Olympio, 1972.
- ROSA, M. C. A. P. *Formação de nomes aumentativos: estudos da produtividade de alguns sufixos portugueses*. Dissertação (Mestrado em Linguística). Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1982.
- SANDMANN, A. J. *Formação de Palavras No Português Brasileiro Contemporâneo*. Curitiba: Ícone. 1989.
- SANTOS, M. J. M. *Os falares fronteiriços de Trás-os-Montes*. Coimbra. Instituto de Língua e Literatura Portuguesas, 1967.
- SEITA, I. F. *A linguagem popular de Aldeia Nova de São Bento*. Dissertação (Licenciatura em Filologia Românica). Faculdade de Letras, Universidade de Lisboa. Lisboa, 1944.
- SILVA, F. J. *Dicionário da língua portuguesa*. 3ª edição, actualizada e valorizada com centenas de locuções latinas, gregas e estrangeiras. Porto: Domingos Barreira, 1955.
- SILVA, M. C. L. R. M. da. Vila Nova de Ourém. *Contribuição para o estudo linguístico, etnográfico e folclórico do concelho*. Dissertação (Licenciatura em Filologia Românica). Faculdade de Letras. Universidade de Coimbra. Coimbra, 1972.
- SILVA, M. G. G. e. Figueiró dos Vinhos. *Contribuição para o estudo da linguagem, etnografia e folclore do concelho*. Dissertação (Licenciatura em Filologia Românica). Faculdade de Letras, Universidade de Coimbra. Coimbra, 1960.
- SOARES BARBOSA, J. *Grammatica Philosophica da Língua Portuguesa ou Princípios da Grammatica Geral Aplicados á Nossa Linguagem*. Lisboa: Typographia Da Acad. Real das Sciencias. S. L., 1822.
- SOUSA, G. S. de. *Tratado descriptivo do Brazil em 1587*. Rio de Janeiro: Typographia Universal de Laemmert, 1851.
- TAVARES, M. C. A toponímia da região centro-norte de Mato Grosso do Sul. In: ISQUERDO, A. N. (org.). *Toponímia ATEMS: caminhos metodológicos*, v.1. Campo Grande, MS : Ed. UFMS, 2019. p. 111-147.
- VITERBO, Fr. J. de S. R. *Elucidário das palavras, termos e frases que em Portugal antigamente se usaram e que hoje regularmente se ignoram. Edição crítica por Mário Fiúza, baseada nos manuscritos do autor*. Porto e Lisboa, Livraria Civilização, vol. I e vol. II. 1965, 1966.